

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2003, 13:08:05 ; Search time 50 Seconds
(without alignments)
2831.682 Million cell updates/sec

Title: CAA47749

Perfect score: 4630

Sequence: 1 MASSPAQRRRRNDPLTSSPG.....LFRMNFSDLRKKMILOOF 892

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4078.5	88.1	904	23	AAO20516 Protein of APP rel
2	4078.5	88.1	904	24	ABR48218 Human bladder canc
3	4078.5	88.1	904	24	ABU56565 Lung cancer-associ
4	2649	57.2	887	22	ABG33343 Drosophila melanog
5	1872	40.4	864	24	ABJ26013 Aspergillus fumiga
6	1840.5	39.8	785	24	ABJ25413 Aspergillus fumiga
7	1808.5	39.1	903	23	ABP73355 Candida albicans e
8	1073.5	23.2	268	21	ABG56883 Human prostate can
9	902.5	19.5	682	22	AAB62036 P. furiosus mcm po

10	891	19.2	1023	24	ABJ26500 Aspergillus fumiga
11	887	19.2	903	24	ABJ26305 Aspergillus fumiga
12	882.5	19.1	831	24	ABJ25705 Aspergillus fumiga
13	875	18.9	167	24	ABU70820 Human adipocyte Se
14	869	18.8	883	22	ABG5203 Human protein sequ
15	860	18.6	874	24	ABJ25900 Aspergillus fumiga
16	846	18.3	910	23	ABP73643 Candida albicans e
17	835	18.0	901	23	ABG66003 P. patens cell cyc
18	834	18.0	878	23	ABF73678 Candida albicans e
19	834	18.0	878	23	ABG93380 C. albicans BAX-as
20	830.5	17.9	880	23	ABF73888 Candida albicans e
21	819	17.7	819	22	ABG60488 Drosophila melanog
22	816.5	17.6	866	22	ABG58299 Drosophila melanog
23	810	17.5	718	24	ABJ25741 Aspergillus fumiga
24	804	17.4	713	24	ABJ26341 Aspergillus fumiga
25	802	17.3	971	23	ABG93369 S. cerevisiae BAX-
26	800	17.3	733	22	ABG60401 Drosophila melanog
27	783	16.9	724	21	ABG56475 Human prostate can
28	778	16.8	766	21	ABG43937 Human cancer assoc
29	776	16.8	921	24	ABJ25714 Aspergillus fumiga
30	774	16.7	720	23	AAO22751 Maize prolifera ge
31	763	16.5	819	23	ABU69420 Lung snail cell ca
32	758.5	16.4	817	22	ABG60368 Drosophila melanog
33	758.5	16.4	840	24	ABU56428 Lung cancer-associ
34	758.5	16.4	841	23	ABG09802 Amino acid sequenc
35	757	16.3	728	23	ABF73710 Candida albicans e
36	754	16.3	957	24	ABJ26314 Aspergillus fumiga
37	753.5	16.3	814	24	ABJ26042 Aspergillus fumiga
38	745.5	16.1	825	23	Amino acid sequenc
39	743	16.0	720	22	ABG61153 Drosophila melanog
40	739.5	16.0	809	23	ABF73372 Candida albicans e
41	732.5	15.8	632	24	ABJ25442 Aspergillus fumiga
42	719	15.5	1049	22	P. furiosus mcm po
43	671	14.5	603	22	AAW94016 Human stomach can
44	671	14.5	603	22	AAE95303 Human protein sequ
45	650.5	14.0	1115	22	AAB96812 Putative P. abyssi

ALIGNMENTS

RESULT 1

AAO20516

ID AAO20516 standard; Protein; 904 AA.

XX AAO20516;

AC AAO20516;

XX 27-JUN-2002 (first entry)

DT Protein of APP related human homologue hCP50592.

DE Neuroprotective; nontropic; transgenic fly; Alzheimer's disease; Abeta;

KW anyloid precursor protein; tissue-specific expression control; human APP;

KW APP pathway modulator; gene therapy.

XX Homo sapiens.

OS WO200226820-A2.

PN 04-APR-2002.

XX 01-OCT-2001; 2001WO-EPI1345.

XX 29-SEP-2000; 2000US-236893P.

PR 14-JUN-2001; 2001US-298309P.

XX (NOVS) NOVARTIS AG.

PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX Cohen D, Dengler UJ, Finelli AL, Freuler F, Konsolaki M;

PI Reinhardt MWHM, Zusman S;

XX WPI; 2002-315796/35.

DR

DR N-PSDB; AAK99410.

XX New transgenic fly, containing DNA encoding an Abeta portion of human

PT APP, useful for identifying agents which modulate the APP pathway and

PT which can be used to treat Alzheimer's disease -

XX Disclosure; Page 123-124; 129pp; English.

PS The invention relates to a transgenic fly whose genome comprises DNA

XX encoding a polypeptide having the Abeta portion of human amyloid

CC precursor protein (APP), fused to a signal sequence. The DNA sequence

CC encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in

CC the specification. The DNA sequence is operably linked to a tissue-

CC specific expression control sequence. Expression of the sequence gives

CC the fly an altered phenotype. The purpose of the invention is for

CC identifying agents that inhibit or promote the expression and/or function

CC of genes or encoded polypeptides which modify the APP pathway. The agent

CC is a compound, triple helix DNA, antisense oligonucleotide, double

CC stranded RNA molecule, ribozyme, or particularly an antibody. It is used

CC to treat conditions such as Alzheimer's disease. The agent can be used as

CC an APP pathway modulator or in gene therapy. This sequence represents the

CC protein of the APP related human homologue hCP50592.

XX

SQ Sequence 904 AA;

Query Match 88.1%; Score 4078.5; DB 23; Length 904;

Best Local Similarity 89.8%; Pred. No. 0;

Matches 817; Conservative 8; Mismatches 52; Indels 33; Gaps 6;

QY 1 MASSPAQRNRNDPLTSPGSRSSRTDALTSPPGRDLPPFDESEGLLGTGCPLEEDBG 60

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 10 MASSPAQRNRNDPLTSPGSRSSRTDALTSPPGRDLPPFDESEGLLGTGCPLEEDBG 69

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 61 EELIGDGMERYRAIPELDAYEAGLALDDDEELTASRREA-----DGPCGVTG 113

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 70 EELIGDGMERYRAIPELDAYEAGLALDDDEELTASRREAERAMRQRDRAGR--- 126

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 114 SWPGIGLCAVGCMTA-----MRTSRALPASASGAGCTGDEDEQMIESTIENLD 165

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 127 ---GLGRMRRLLYDSDEERPAKRQRVERA-----TEDGEDEEMISTENLED 176

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 166 LKSHSVRWVSMAGPRLEIHRFNFLRTHVDSHGHNFKERISDMCKENRESLVNYED 225

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 177 LKSHSVRWVSMAGPRLEIHRFNFLRTHVDSHGHNFKERISDMCKENRESLVNYED 236

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 226 LAAREHVLAYFLPAPALLOIFDEAALEVVLAMYPKYDRITNHNHVRISHLPLVEELRS 285

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 237 LAAREHVLAYFLPAPALLOIFDEAALEVVLAMYPKYDRITNHNHVRISHLPLVEELRS 296

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 286 LRQLHLNQLIRTSVVTSGTGLVLPQLSMVKYCNKCNFVLGPFQSQNQEVKPGSCPEQ 345

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 297 LRQLHLNQLIRTSVVTSGTGLVLPQLSMVKYCNKCNFVLGPFQSQNQEVKPGSCPEQ 356

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 346 SAGPFEVNMETIYQNYQRIQRIQSPGKVAARRLPRSKDAILLADLVDSNAGDEIELTG 405

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 357 SAGPFEVNMETIYQNYQRIQRIQSPGKVAARRLPRSKDAILLADLVDSNAGDEIELTG 416

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 406 IYHNYDGLNTANGFPVFATVILANHVAKKDNKAVAGELTDEDVKMTITSLSKDQIGEK 465

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 417 IYHNYDGLNTANGFPVFATVILANHVAKKDNKAVAGELTDEDVKMTITSLSKDQIGEK 476

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 466 IFASTAPSIYGHEDIKRGPAALFEGEPKPGKHVKGVDINVLCCGDPGTAKSQFYKI 525

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 477 IFASTAPSIYGHEDIKRGPAALFEGEPKPGKHVKGVDINVLCCGDPGTAKSQFYKI 536

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 526 EKVSRAIFTTGQASAVATYVQRHPVSRWLEAGALVLRGVCLIDFPMKNDQ 585

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 537 EKVSRAIFTTGQASAVATYVQRHPVSRWLEAGALVLRGVCLIDFPMKNDQ 596

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 586 RTSHEAMEQOISISKAGIYVTSIQARCTVIAAANPIGGRYDPSLTFSENVDLTEPIISR 645

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 597 RTSHEAMEQOISISKAGIYVTSIQARCTVIAAANPIGGRYDPSLTFSENVDLTEPIISR 656

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 646 FDILCVVRDVTDPVQDEMLARFVGVSHVRHHPSNKEEGLANGSAAEPAMPNTYGVPEPLP 705

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 657 FDILCVVRDVTDPVQDEMLARFVGVSHVRHHPSNKEEGLANGSAAEPAMPNTYGVPEPLP 716

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 706 QEVLLKKYIIYAKERVHPKLNQDQDKVAKMYSDLRKESMATGSPITVVRHIESMHSHGCGP 765

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 717 QEVLLKKYIIYAKERVHPKLNQDQDKVAKMYSDLRKESMATGSPITVVRHIESMIRMA-- 774

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 766 RAHPSAGLCLDRRRRQHGHPDRAGELHRHTEYQRH---RSMRKTFFARYLSFRDNNELLFL 822

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 775 EAHARIHLRDVIEDDYNMAIRVMELESFIDTQKESVMSMRKTFFARYLSFRDNNELLFL 834

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 823 ILKOLVABQVYQYQNRFGAQQDTIEVPEKDLVDKARQINIHLSAFYDSELFMRMKNFSDH 882

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 835 ILKOLVABQVYQYQNRFGAQQDTIEVPEKDLVDKARQINIHLSAFYDSELFMRMKNFSDH 894

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 883 LKRRKMILQOF 892

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 895 LKRRKMILQOF 904

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2

ABR48218

ID ABR48218 standard; Protein: 904 AA.

XX

AC ABR48218;

XX

DT 12-JUN-2003 (first entry)

XX

DE Human bladder cancer associated protein sequence SEQ ID NO:155.

XX

KW Human; bladder cancer; cytostatic; gene therapy; vaccine.

XX

OS Homo sapiens.

XX

PN WO2003003906-A2.

XX

PD 16-JAN-2003.

XX

PF 03-JUL-2002; 2003WO-US21338.

XX

PR 03-JUL-2001; 2001US-302814P.

XX

PR 03-AUG-2001; 2001US-310099P.

PR 08-NOV-2001; 2001US-343705P.

PR 13-NOV-2001; 2001US-350666P.

PR 12-APR-2002; 2002US-372246P.

XX

PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX

PI Mack DH, Aziz N;

XX

WI: 2003-201532/19.

DR N-PSDB; ACC51032.

XX

PT Detecting a bladder cancer-associated transcript in a cell from a

PT patient, comprises contacting a biological sample from the patient with

PT a bladder cancer-associated polynucleotide or antibody -

XX

PS Claim 10; Page 283-284; 307pp; English.

XX

CC The present invention describes a method for detecting a bladder cancer-

CC associated transcript in a cell from a patient. The method comprises

CC contacting a biological sample from the patient with a polynucleotide

CC that selectively hybridizes to a sequence that is 80 % identical to a

CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059

CC encode the human bladder cancer-associated proteins given in ABR48146 to

CC ABR48242). Bladder cancer-associated sequences from the present invention

CC have cytostatic activities, and can be used in antisense gene therapy and

CC in vaccine production. The method can be used for detecting a bladder

CC cancer-associated transcript in a cell from a patient. The method is

CC useful in diagnosing or treating bladder cancer and in screening for

CC compounds that modulate bladder cancer, such as hormones or antibodies.

CC The nucleic acid molecules from the present invention may be used in

CC various screening and diagnostic methods, and for gene therapy, vaccine
XX and/or antisense/inhibition applications.
SQ Sequence 904 AA;
Query Match 88.1%; Score 4078.5; DB 24; Length 904;
Best Local Similarity 89.8%; Pred. No. 0;
Matches 817; Conservative 8; Mismatches 52; Indels 33; Gaps 6;
QY 1 MASSPAQRRRGNPLTSSPGRRSRRTDALTSSPGRLDPPEDESEGLLTGTEPLEEEDG 60
DB |||||
QY 10 MASSPAQRRRGNPLTSSPGRRSRRTDALTSSPGRLDPPEDESEGLLTGTEPLEEEDG 69
DB |||||
QY 61 EELIGDMERYAIRPELDAYEAGLALDDEDEVEELTASREAA-----DPCGCVTG 113
DB |||||
QY 70 EELIGDMERYAIRPELDAYEAGLALDDEDEVEELTASREAAERAMRQDREAGR--- 126
DB |||||
QY 114 SWFGLGACAVGSCWTA-----MRTSRALPASASAGAGTDEGEDEQMIESIENLED 165
DB ||||| : : : : :
QY 127 ---GLGRMRGLLYDSEDEERPAKRQVERA-----TEDGEDEQMIESIENLED 176
DB ||||| : : : : :
QY 166 LKGHSVREWSMAGPLRIHHRKFNFLRTHVDSGHNVFKERISDMCKENRESILVNYED 225
DB |||||
QY 177 LKGHSVREWSMAGPLRIHHRKFNFLRTHVDSGHNVFKERISDMCKENRESILVNYED 236
DB |||||
QY 226 LAAREHVLAYFLPEAPAEALQIFDEAALEVVLAMYPKYDRITNHHIVRISHLPLVERLS 285
DB |||||
QY 237 LAAREHVLAYFLPEAPAEALQIFDEAALEVVLAMYPKYDRITNHHIVRISHLPLVERLS 296
DB |||||
QY 286 LRQLHLNQLIRTSQVTSCTGVLPOLSMVYKNCNKFVLGPFQCSQONQEVKPGSCPECQ 345
DB |||||
QY 297 LRQLHLNQLIRTSQVTSCTGVLPOLSMVYKNCNKFVLGPFQCSQONQEVKPGSCPECQ 356
DB |||||
QY 346 SAGPFEVNMETIYQNTORIQESPCGKVAARLPRSKDAILLADLVDSNAGDELTG 405
DB |||||
QY 357 SAGPFEVNMETIYQNTORIQESPCGKVAARLPRSKDAILLADLVDSNAGDELTG 416
DB |||||
QY 406 IYHNNDGSLNTANGFPVATVILANHVAKKDNKVAAGELTDEDVKMITSLSKDDQIGEK 465
DB |||||
QY 417 IYHNNDGSLNTANGFPVATVILANHVAKKDNKVAAGELTDEDVKMITSLSKDDQIGEK 476
DB |||||
QY 466 IFASIAPSIYGHEDIKRGPAALFPGGPKNGKHKVKGDIINVLGDPGTAKSQFLKYI 525
DB |||||
QY 477 IFASIAPSIYGHEDIKRGPAALFPGGPKNGKHKVKGDIINVLGDPGTAKSQFLKYI 536
DB |||||
QY 526 EKVSSRAIFTGGASAVATVYVORHPVSRWLEAGALVADRGVCLTDEFKMDQD 585
DB |||||
QY 537 EKVSSRAIFTGGASAVATVYVORHPVSRWLEAGALVADRGVCLTDEFKMDQD 596
DB |||||
QY 586 RTSIHEAMEOOSISIKAGIVTSLOARCTVIAAANPIGGRYDPSLTFSENVDLTEPTISR 645
DB |||||
QY 646 FDILCVVRDVPVQDEMLARFVGVSHRHPNSKEEGLANGSAAPAMPNTYGVPEPLP 705
DB |||||
QY 657 FDILCVVRDVPVQDEMLARFVGVSHRHPNSKEEGLANGSAAPAMPNTYGVPEPLP 716
DB |||||
QY 706 QEVLKVIYKAKERVHPKLNQMDQDKVAKMYSDLRKESMATGSPITVRIHESMHHGGP 765
DB |||||
QY 717 QEVLKVIYKAKERVHPKLNQMDQDKVAKMYSDLRKESMATGSPITVRIHESMIRNA-- 774
DB ||||| : : : : :
QY 766 RAHPSAGLDRRRRQHGHPDAGELHRRHTEVQRH---RSMRKTFAFYLSPFRDNNELELLF 822
DB ||||| : : : : :
QY 775 EAHARIHLRVYIEDDVMNAIRVMLESFIDTQFVMSRMKKTFAFYLSPFRDNNELELLF 834
DB ||||| : : : : :
QY 823 ILKQVLAQVYTYORNRFGAQQDTIEVPEKDLVDKARQINIHNLSAFYDSELFRMKNFSD 882
DB |||||
QY 835 ILKQVLAQVYTYORNRFGAQQDTIEVPEKDLVDKARQINIHNLSAFYDSELFRMKNFSD 894
DB |||||
QY 883 LKRKMILQQF 892
DB |||||
QY 895 LKRKMILQQF 904

RESULT 3
ABU56565
ID ABU56565 standard; Protein; 904 AA.
XX AC ABU56565;
XX DT 02-APR-2003 (first entry)
XX Lung cancer-associated polypeptide #158.
XX Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX Unidentified.
OS WO200286443-A2.
PN 31-OCT-2002.
XX 18-APR-2002; 2002WO-US12476.
XX 18-APR-2001; 2001US-284770P.
PR 10-MAY-2001; 2001US-290492P.
PR 09-NOV-2001; 2001US-339245P.
PR 13-NOV-2001; 2001US-350666P.
PR 29-NOV-2001; 2001US-334370P.
PR 12-APR-2002; 2002US-372246P.
XX (E0SB-) EOS BIOTECHNOLOGY INC.
XX Aziz N, Murray R;
PI WPI; 2003-093161/08.
XX N-PSDB; ABX76294.
DR Detecting a lung cancer-associated transcript in a cell from a patient
DR for treating lung cancer, by contacting a biological sample from the
DR patient with a polynucleotide that exhibits increased or decreased
XX expression in lung cancer
PS Claim 27; Page 310; 453pp; English.
XX The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung
CC cancer-associated polynucleotides and polypeptides are used for
CC identifying a compound that modulates a lung cancer-associated
CC polypeptide, for inhibiting proliferation of a lung cancer-associated
CC cell to treat lung cancer in a patient and for treating a mammal having
CC lung cancer by administering a modulatory compound identified. The
CC methods are useful for treating lung cancer, such as small cell lung
CC cancer, non-small cell lung cancer or other benign or precancerous
CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
CC and polypeptides are useful for diagnostic purposes and as targets for
CC screening for therapeutic compounds that modulate lung cancer, such as
CC antibodies. Sequences ABU56408-ABU56745 represent lung cancer-associated
CC polypeptides of the invention.
SQ Sequence 904 AA;
Query Match 88.1%; Score 4078.5; DB 24; Length 904;
Best Local Similarity 89.8%; Pred. No. 0;
Matches 817; Conservative 8; Mismatches 52; Indels 33; Gaps 6;

QY 1 MASSPAQRRRGNPLTSSPGRRSRRTDALTSSPGRLDPPEDESEGLLTGTEPLEEEDG 60

Qy	405	GIYHNNDGSLNTANGFPVFATVILANHAVAKKNKNAVGLTDBDVKMITSLSKDQOIG	464
Db	401	GIYTNNDGSLNTDQGFVFATVIAHNVVYKDSKVQVSLTDEDIATIQKLSKDPRIE	460
Qy	465	KIFASTAPSIYGHEDIKRGPALALFGGEPKNPGCKHKVRGDIINVLLCGDGPCTAKSOLKY	524
Db	461	RVVASMAPSIYGHDIYIKRALALALFGGESKNPGCKHKVRGDIINLLICGDGPCTAKSOLKY	520
Qy	525	IEKVSRAIFTTTGGQASAVATVYQVRHPVSRWTLFAGALVLDARGVCLIDFDMKNDQ	584
Db	521	TEKVAPRAVFTTGGQASAVGLTAVVRNPVSRWTLFAGALVLDAGVCLIDFDMKNDQ	580
Qy	585	DRTSIHEAMEQOSISISKAGIVTSIQARCTVIAAANPIGGRYDPSLTFSENVDLTPEIS	644
Db	581	DRTSIHEAMEQOSISISKAGIVTSIQARCTVIAAANPIGGRYDPSMTFSENVLSEPLS	640
Qy	645	RDILLCVVRDTPVDQDEMLARFVVGSHVRHHPNKEEGLANGSRAEPAWPNTYGVPL	704
Db	641	RDVLCVVKDEDFMQDQDLAKFVHSHMKHHPSEEQPELE-----PQLKTVDEI	692
Qy	705	POEVLKYYIIYAKERVHPKLNQMDQKVAKMYSLDKRESMATGISPITVRHIESMSHGG	764
Db	693	QDQLLRQXIVYAKENIRPKLTNIDEDKIAKWAQLRQESFATGSLPTIVRHIESVIRMS-	751
Qy	765	PTAHSAGLCDDRRRHQGHPRDAG-----ELHRTTEVORH---RSMRKTTFARYLSFRDN	816
Db	752	-BAHARMHL-----RENYMEADVSNMIRMLESTEAQKFSVMKKMRSTFKYLSFKDH	805
Qy	817	NELLFILKQLVAEOVTVYORNFQAQODTIEVPEKLDVKARQININLSAFYDSELFPM	876
Db	806	SELLFILKQLTDLQALYIRCKDGGATHVEIMERDLIERAKQLDIVNLKAPFYESDLFT	865
Qy	877	NKFSHDLKRKMILQ 890	
Db	866	NGFSYDPKRRILQ 879	
RESULT 5			
ABJ26013			
ID	ABJ26013	standard; Protein; 864 aa.	
XX	ABJ26013;		
XX			
DT	16-APR-2003	(first entry)	
XX	Aspergillus fumigatus essential gene protein #671.		
DE			
XX			
KW	Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;		
KW	cancer; contamination; biofilm; antibody; immune response.		
XX			
OS	Aspergillus fumigatus.		
XX			
PN	WO200286090-A2.		
XX			
PD	31-OCT-2002.		
XX			
PF	23-APR-2002; 2002WO-US13142.		
XX			
PR	23-APR-2001; 2001US-285697P.		
PR	27-APR-2001; 2001US-287066P.		
PR	05-JUN-2001; 2001US-295899P.		
PR	09-JUL-2001; 2001US-303899P.		
PR	31-AUG-2001; 2001US-316362P.		
XX			
PA	(ELIT-) ELITRA PHARM INC.		
XX			
PI	Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SW;		
XX			
DR	WPI; 2003-093124/08.		
XX			
PT	New purified or isolated nucleic acids of essential genes of		
PT	Aspergillus fumigatus, useful for treating or preventing infections by		

QY 742 ESMATGSIPTV 753
II:III: IIII
Db 761 ESLATGAYPTV 772

RESULT 7
ABP73355
ID ABP73355 standard; Protein; 903 AA.
XX AC ABP73355;
XX 30-JAN-2003 (first entry)
XX Candida albicans essential protein SEQ ID NO 7192.
XX Funus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
KW signal transduction; DNA replication; cell division; growth;
KW proliferation; Candida albicans; fungicide; antifungal.
XX
OS Candida albicans.
XX WO200253728-A2.
XX
XX 11-JUL-2002.
XX
XX 26-DEC-2001; 2001WO-US49486.
XX
XX 29-DEC-2000; 2000US-259128P.
XX 20-FEB-2001; 2001US-0792024.
XX 22-AUG-2001; 2001US-314050P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
XX WPI; 2002-566694/60.
XX N-PSDB; AB231905.
XX
XX Constructing strains for identifying gene products as effective targets
XX for therapeutic intervention, by inactivating in the strain one allele
XX of a gene and placing other allele of the gene under conditional
XX expression -
XX
XX Claim 44; SEQ ID NO 7192; 167pp + Sequence Listing; English.
XX
XX The invention relates to constructing (M1) a strain of diploid fungal
XX cells in which both alleles of a gene are modified, comprising modifying
XX one allele by insertion or replacement by a cassette having an
XX expressible selectable marker and modifying other allele by
XX recombination, of a promoter replacement fragment with a heterologous
XX promoter, so that expression of the second allele is regulated by the
XX promoter. (M1) is useful for constructing a strain of diploid fungal
XX cells in which both alleles of a gene are modified. The diploid fungal
XX cells having both alleles modified are useful for identifying a gene that
XX is essential to the survival or growth of a fungus, a gene that
XX contributes to the virulence and/or pathogenicity of a fungus, a gene
XX that contributes to the resistance of a diploid fungus to an antifungal
XX agent, an antifungal agent that inhibits the growth of a diploid fungus
XX and for identifying a therapeutic agent for treatment of a mammalian
XX disease. (M1) is useful for identifying a compound which modulates the
XX activity of a gene product, preferably enzymatic activity, carbon
XX compound catabolism, biosynthetic, transporter, transcriptional,
XX translational, signal transduction, DNA replication and cell division
XX activity. The method is useful for identifying a compound having the
XX ability to inhibit growth or proliferation of C. albicans cells and for
XX treating infection by C. albicans. The present sequence is that of an
XX essential Candida albicans protein used in the method of the invention.
XX Note: The sequence data for this patent is not represented in the printed
XX specification but is based on sequence information supplied to Derwent by
XX the European Patent Office.

Sequence 903 AA;

Query Match 39.1%; Score 1808.5; DB 23; Length 903;
Best Local Similarity 45.0%; Pred. No. 2e-153;
Matches 393; Conservative 124; Mismatches 216; Indels 139; Gaps 14;

QY 1 MASSPA-----QRRRGNDPL-----TSSP--GRSSRRRTDALTSSPG 34
II:III: IIII
Db 1 MSSPPAASSSRDNNRRKRSPYDDEDDDDNNNDNRSSIHQSPSSPGLGSPNPHQLPSSPA 60
II:III: IIII
QY 35 -----RDLPPFESEGLGTGEGPLEEEDGEEELIGDGMEDYRA 74
II:III: IIII
Db 61 IPFDALDVEDVEEIVHNDIQDLNFPSPS-----EEEDDLM-EMLEQDYRA 107
II:III: IIII
QY 75 IPELDAYE-AEGLALDDEDEVELTASREAAADGPGCTVTGSPGLCACAVGSCMTAMRRT 133
II:III: IIII
Db 108 NFEQDHYLDGNDIDTGDYEEMDAATR-----RMINQWNRDAI-----LNNANRSR 156
II:III: IIII
QY 134 RSA-----LPASASAGTDEGEDEEQMTE-----SIE 161
II:III: IIII
Db 157 RGAFLDDDDDEDEDEDESMGQFGLPIORRRRRRHHDEDDQDDMDVDVETDPTNEELSLE 216
II:III: IIII
QY 162 NLEDLKGHSVRWVSMAGPRLEIHHRFKNFLRTHVDHSHGNVFKERISDMCKENRESLVV 221
II:III: IIII
Db 217 SUSDKAPSITEWILQPAVSRSTARELKSFFLEYTDANGDSVYGNKMRTLGEVNAESLEV 276
II:III: IIII
QY 222 NYEDLAAREHVLAYFLPEAPAEALLOIFDEAALEVLAMYPKYDRITNHHIVRISHLPLVE 281
II:III: IIII
Db 277 SYKDLADSKAILALFLATSPHEMLKIFDIVAMEAVELHYPNYSQIHQEVHVRTDFPNIL 336
II:III: IIII
QY 282 ELRSURLHLNQLIRTSVVTSGTVLPQLSVMKYNCNKNFVLGPFQCSQNVKVGKSC 341
II:III: IIII
Db 337 NLRDLRESNLQLVKVGVVTRTGVFIQLKYVKFDCLKCGVVLGVPYVODSNTEVKSIFC 396
II:III: IIII
QY 342 PCQSAGPFEVNMETIYQYRIQESPGKVAARLPRSKDAIILLADLVDSNAGDEI 401
II:III: IIII
Db 397 TNCQSGKPGPKLNSEKTLRYNYRITLQEAQPTVPAGRLPHRHRVILLSDLVDVAKPGEDI 456
II:III: IIII
QY 402 ELTGIYHNNYDGLSNTANGFPVFATVILANHAKDKNVAVG-----ELTDHDDVKMITS 455
II:III: IIII
Db 457 EVTGIYKNYDGLNNAKNGFPVFATILEANSIRKESSAFMGNNLVNMMWTEIREFRK 516
II:III: IIII
QY 456 LSKDOOIGKIFASTAPSIYGHEDIKRGPALALFGGEPKNPGKHVKRGDINVLLCGDPG 515
II:III: IIII
Db 517 LSKEGIIDKIIASMAPSIYGHDKIKTATACSLFGGVPKDVNGKLSIRGDIVNLLGDPG 576
II:III: IIII
QY 516 TAKSOFKLYIEKVSSRAIETTGOGASAVAVTAYVORHPVSRWETLEAGALVLADRGVCLI 575
II:III: IIII
Db 577 TAKSQLKYAEKTASRAVATGOGASAVGLTASVRKDPITRETLEGGALVLADRGVCLI 636
II:III: IIII
QY 576 DEFDKMNDDRTSIHEAMEQQSISISKAGIVTSIQARCTVIAAANPIGRYDP-SLTFSEN 635
II:III: IIII
Db 637 DEFDKMNDDRTSIHEAMEQQSISISKAGIVTTHARCAVIAAANPNNGRYNSTLPLPEN 696
II:III: IIII
QY 636 VDLTEPISRFDILCVVRDTPVDQDEMLARFVGVSHVRHHPNKEEGLANGSAAPPA- 694
II:III: IIII
Db 697 VDLTYPILSRFDIMCIVRDLVNPESDERLASFVIDSHMRSHPTNEEDILNANGKGSQAQ 756
II:III: IIII
QY 695 -----MPNYGYVE-----PLPOEVKLYIYAKERVHPKLN 725
II:III: IIII
Db 757 DDENNMEDDGVDPQPSAARTSERIEQLNKEQEETSPIPQDILLIKYIQIYARVKIOPKHL 816
II:III: IIII
QY 726 QMDQKQVAKVYSDLRKESMATGSIPTVTRHIESM 759
II:III: IIII
Db 817 QMNMKLVARYADLRKEATTCGSPITVRHLESI 850
II:III: IIII

RESULT 8
AAB56883
ID AAB56883 standard; Protein; 268 AA.
XX AC AAB56883;
XX
XX 13-MAR-2001 (first entry)

```
XX Human prostate cancer antigen protein sequence SEQ ID NO:1461.
DE
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytoskeletal; cardioactive; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotropic; antineoplastic; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.
XX
OS Homo sapiens.
XX
XX WO200005174-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05988.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2000-587513/55.
XX
XX N-PSDB; AAF16086.
XX
XX Prostate cancer associated gene sequences, referred to as prostate
XX cancer antigens, useful for treatment, prevention, and diagnosis of
XX disorders such as prostate cancer -
XX
XX Claim 11; Page 1894-1895; 2338pp; English.
XX
XX AAF15566 to AAF16505 encode the human prostate cancer associated
XX proteins, called prostate cancer antigens, given in AAF56363 to AAF57302.
XX The prostate cancer antigens can have neuroprotective, cytostatic,
XX cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
XX nephrotropic, antineoplastic, gynaecological and antibacterial activities,
XX and can be used in gene therapy. The prostate cancer antigen
XX polynucleotides may be used for detection of prostate cancer, chromosome
XX identification, as chromosome markers, and for numerous other diagnostic
XX or research purposes. The prostate cancer antigens may be used to treat
XX disorders such as neural, immune, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX AAF57303 represent sequences used in the exemplification of the present
XX invention.
XX
XX Sequence 268 AA;
XX
XX Query Match 23.2%; Score 1073.5; DB 21; Length 268;
XX Best Local Similarity 91.7%; Pred. No. 5.8e-88;
XX Matches 211; Conservative 5; Mismatches 13; Indels 1; Gaps 1;
XX
XX 1 MASSPAQRNRNDPLTSSPGRRSRRTDALTSSPGRDLPPEDESEGLLGTGEPLEEDG 60
XX
XX 30 MASSPAQRNRNDPLTSSPGRRSRRTDALTSSPGRDLPPEDESEGLLGTGEPLEEDG 89
XX
XX 61 BELIGDGMERYDRAIPELDYEAEGLALDDDEVEELTASREAAADGCGVTGSGWFL-G 119
XX
XX 90 BELIGDGMERYDRAIPELDYEAEGLALDDDEVEELTASREAAAMRHVTGRLAGAWA 149
XX
XX 120 ACAGVSGMTAMRTSRPALPASASAGCTGDEGEDEQMIENLEDLKGHSVREWSMAG 179
XX
XX 150 ACAGVSGMTAMRTSRPALPASASAGCTGDEGEDEQMIENLEDLKGHSVREWSMAG 209
XX
XX 180 PRLEIHRFRKFNFLRTHVDSHGHNFKERISDMCKENRESLVVNYEDLAAR 229
XX
XX 210 PRLEIHRFRKFNFLRTHVDSHGHNFKERISDMCKENRESLVVNYEDTGSQ 259
XX
XX RESULT 9
```

```
AAB62036
ID AAB62036 standard; Protein; 682 AA.
XX
AC AAB62036;
XX
DT 14-MAY-2001 (first entry)
XX
DE P. furiosus mcm polypeptide with intein absent.
XX
KW MCM; minichromosome maintenance protein; archaeal polypeptide; PCNA;
KW RFC-P38; RFC-P55; RPA; CDC6; FEN-1; dUTPase; ligase; helicase dna2;
KW nucleic acid amplification; polymerase chain reaction; PCR.
XX
OS Pyrococcus furiosus.
XX
XX WO200109347-A2.
XX
XX 08-FEB-2001.
XX
XX 28-JUL-2000; 2000WO-US20532.
XX
XX 30-JUL-1999; 99US-0146580.
XX
XX (STRA-) STRATAGENE.
XX
XX Hogrefe HH, Cline JM, Hansen CJ, Borns MC;
XX
XX WPI; 2001-182959/18.
XX N-PSDB; AAF57040.
XX
XX Composition for improving nucleic acid polymerase reactions, useful
XX e.g. in synthesis or amplification, contains at least one archaeal
XX accessory protein -
XX
XX Claim 189; Fig 47; 147pp; English.
XX
XX The invention provides a composition (A) for enhancing nucleic acid
XX polymerase reactions that comprises an archaeal MCM (minichromosome
XX maintenance protein) and at least one of the archaeal polypeptides (PCNA,
XX RFC-P38 or -P55, RPA, CDC6, FEN-1, dUTPase, ligase, helicase dna2, or
XX helicases 2-8). (A) And similar compositions containing different
XX combinations of accessory proteins, are used to improve performance of
XX synthesis, amplification, mutagenizing, labeling and detecting reactions,
XX e.g. for gene characterization, cloning, detection of allelic variants,
XX diagnosis and screening for disease, particularly where done by
XX polymerase chain reaction (PCR). Some of the proteins also stabilize
XX duplexes during polymerase reactions or improve exonuclease reactions,
XX for example RFA also improves specificity of nucleic acid/protein
XX interaction and PCNA improves polymerase-mediated repair processes and
XX hybridization reactions. Nucleic acids encoding the archaeal polypeptides
XX are used for recombinant production of proteins, and fragments of the
XX nucleic acid as probes and primers for screening related sequences. The
XX accessory proteins increase accuracy and efficiency of polymerase
XX reactions, allow use of lower denaturation and extension temperatures
XX (possibly isothermal processing), and improve synthesis of long targets.
XX The present sequence represents the amino acid sequence of P. furiosus
XX mcm polypeptide with intein absent.
XX
XX Sequence 682 AA;
XX
XX Query Match 19.5%; Score 902.5; DB 22; Length 682;
XX Best Local Similarity 34.5%; Pred. No. 7.3e-72;
XX Matches 243; Conservative 137; Mismatches 254; Indels 71; Gaps 21;
XX
XX 181 RLEIHRFRKFNFLRTHVDSHGHNFKERISDMCK-ENRESLVVNYEDLAAREHVLAYFLPE 239
XX
XX 3 REMIERFANFLREYDTEDEGNPVIRGKITDILLITPRSRVAIDMHLNSFDSSELAHEVIE 62
XX
XX 240 APAELLIQIFDEAALEVVLMAPYKDYRITNHIHVRISHLPLVEELSRQLHNLQIRTSQ 299
XX
XX 63 NPEEGISAAEDAQIVLRDFQRED--VGKHARFYNLPETLMVKDGAEHINKLIQVEG 120
XX
XX 300 VVTSTCTGVLPLQSLMWKYNCKNCF-VLGPFCQSQNQVKPGSCPEQSGAGPFVNMMEETI 358
```

```
Db 121 IYTRVGEIKPFVSVAVFVCKDCHEMIVPQKPYESLE-KVKCEQCGSKN-IELDYNKSS 178
QY 359 YONYQIRIQESPGKVAARRPRSKDAILLADLVDSNAGDETELGTGIYHNNYDGLNTA 418
Db 179 FYNQSFRLQDRPETLKGEMPRFDIGILLDDIVDALPGDRVIVTGIILRVLEKREKT- 237
QY 419 NGFPVATVILANHVAKDNKVAAGBELTDEKMTSLSKDQIGIKFASIPASTYIGHE 478
Db 238 --PIPRKILEVNHIEPVSKQEQLSPSEEQIIEKELAKRDKDIVDAIVDSIAPAIYGVK 294
QY 479 DIKRGPALALFGG-EPKNPGGKHKVRGDIINVLGDPGTAKSFLYIEKVSRAIFTTG 537
Db 295 EVKKGITATLALFGGVSRLKPDGT-RLRGDIHVLVGGDPGVAKSQILRYVANLAPRAIYTS 353
QY 538 QGASAVATYAYQVRHPVPSREWLEALGALVLAGDVGCLIDFEDKMNDDRTSIHEAMEQOS 597
Db 354 KSSAAGLTAARVDEFTGGWLEAGALVLAGDYGALIDELDKMSDRDSVHEALEQQT 413
QY 598 ISISKAGIVTSLOARCTVIAAANPIGGRYDPSLTFSENVDLTPFIISREFDILCVVRD 657
Db 414 ISISKAGITATLNARTVTIAAANPKOGRNRMKNPPEQIDLPTLLSRFDLIFVLIDEP 473
QY 658 PVODEMLARFVGVSHVRHPSKEEGLANGSAEPAEMNTYGVPELPDOELVKYIYAK 717
Db 474 DKIDSEVARHIL--RVRRGES-----EVVAPK-----IPHEILKRYIAYAR 512
QY 718 ERVHPKLNOMDQKVAKMSDLRK-----BSMATGSIPTVVRHTESMHHGGPRAHPS 770
Db 513 KNIHPVISEEMEEIEKYIVRMKSKVKYKGEEGIPPIPTARQLEALIRLS--BAHAR 570
QY 771 AGLCRRRRQHGHPDAGELHHRHTEVQRHRSRKTT-----FARYLSFRDRNNE 818
Db 571 MRLSPVITREDA--REAIKMEY--LKOIANDETGQIDVTILELQOSARKLS----KIE 622
QY 819 LLLFILKQLVAEOVTVQRNFGAQ-ODTIEVPEKOLVDK--ARQI 860
Db 623 KILDIIEKL---QKTSER---GAHVNDIILEAKKAGIERQAREI 661

RESULT 10
ABJ26500
ID ABJ26500 standard; Protein; 1023 AA.
XX
AC ABJ26500;
XX
DT 16-APR-2003 (first entry)
XX
DE Aspergillus fumigatus essential gene protein #1158.
XX
KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
KW cancer; contamination; biofilm; antibody; immune response.
XX
OS Aspergillus fumigatus.
XX
PN WO200286090-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US13142.
XX
PR 23-APR-2001; 2001US-285697P.
PR 27-APR-2001; 2001US-287066P.
PR 05-JUN-2001; 2001US-295890P.
PR 09-JUL-2001; 2001US-303899P.
PR 31-AUG-2001; 2001US-316362P.
XX
PA (ELITR) ELITRA PHARM INC.
XX
PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX
DR WPI; 2003-093124/08.
XX
```

```
PT New purified or isolated nucleic acids of essential genes of
PT Aspergillus fumigatus, useful for treating or preventing infections by
PT A. fumigatus, or for treating a non-infectious disease in a subject
XX e.g. cancer
PS Disclosure; Page -: 175pp; English.
XX
CC The invention relates to novel purified or isolated nucleic acids of
CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
CC the invention are used to treat or prevent infections by a pathogenic
CC organism such as A. fumigatus, to treat a non-infectious disease in a
CC subject (e.g. cancer), to prevent or contain contamination of an object
CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
CC biofilm comprising A. fumigatus. The polynucleotides are useful for
CC expressing recombinant protein for characterisation, screening or
CC therapeutic use, as markers for host tissues in which the pathogenic
CC organisms invade or reside, for comparing with the DNA sequence of A.
CC fumigatus to identify duplicated genes or paralogues having the same or
CC similar biochemical activity and/or function, for comparing with DNA
CC sequences of other related or distant pathogenic organisms to identify
CC potential orthologous essential or virulence genes, for selecting and
CC making oligomers for attachment to a nucleic acid array for examination
CC of expression patterns, for raising anti-protein antibodies, as an
CC antigen to raise anti-DNA antibodies or to elicit another immune
CC response, and for identifying polynucleotides encoding the other protein
CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case of virulence
CC factors. This sequence represents a protein of one of the essential genes
CC of Aspergillus fumigatus of the invention.
XX
SQ Sequence 1023 AA;
```

Query Match 19.2%; Score 891; DB 24; Length 1023;

Best Local Similarity 28.7%; Pred. No. 1.6e-70;

Matches 268; Conservative 140; Mismatches 304; Indels 222; Gaps 29;

```
QY 1 MASSPAQRBRG---NDPLTSSPGSRSSRRRTDALTSSPGRDLPPEFE-SEGLLTGEPLEE 56
Db 1 MSSPASSRRRGPRAKDTATSSPARSTRKSLQLOTSSP---TPRAADSOQATPRASRLR- 56
QY 57 EEDGEELIGDMERDYRAIPELDAYEAGLALDDEVEDLTASRRRAAD-----GP 107
Db 57 ---GEAAPSSSPMFFQSSPSKADSSAE---TPDVRMDEPSSPMRESSTWDEGDRPRGN 110
QY 108 CGTVTGWPGLGACAVGSCMTAMRR--TRSALPASA-----141
Db 111 APTMRDSSP-IRYMSSSSPTRAQNRQRSSRDIPSSSSGLFVSRSPSTESNRAVSRRSDLH 169
QY 142 -----ASCAGTDEGEDEQMIESTENL-----EDLKGHSVRE-WVS 176
Db 170 SGGFLSSPNRRRVFDANGMPATDG--DPRSDATFSNIHPDTSEAPALGSGSTRVW-- 225
QY 177 MAGPRLIEH---HREKNFL-----RTHVDSHGHNVFKERTSDCKENR----- 217
Db 226 --GTNISIQDSMSAFKNFLYNFQTKYRLWAEGATEDETRIMGDSAEEREYISMLSTMRL 283
QY 218 ---SLVVNYEDLAAREHVLAYF--LPEAPAEALQIFDEALEVVLAM-----259
Db 284 GVTSLNDAKNAKAYPSTLKLWHLHAYPOEIIPLMDQTVKDVWVELAIKEMERLRAQNO 343
QY 260 -----YPKYDRITNHIHVRISHLP-LVEE-----LRSL 286
Db 344 RNQHNHRLGSSGPAVPSDALSSETGRMPQNEIPDLVGEVETKAFKVLPGFLDSTVNNRDL 403
QY 287 RQLHLNOLIRTSQVWTSCGTGVLPSLMVKYKNCNKFVLCGPFQCSQNOEQKPCSCP--C 344
Db 404 DPADMOKLVSIKGLVIRTPPIIDMKAEAFRCQVCHGV-QVDIDRCKIAPETCEPRPVC 462
QY 345 QSAGPPEVNMEETIYQYQIRIQESPGKVAARRLPRSKDAILLADLVDSNAGDEIELT 404
```



```

Db      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
463 KERNMQLHNRVCFADQVQIKLQETPDSIPDQGTPHSVSLCVYDELVDVCKAGDRVEVT 522
Qy      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
405 GIYHNNYDGLSANTGFPV-----FATVILANHVAKKDNK-----439
Db      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
523 GIFRCN-----PVRNPRQTKQSLFTYIDVLHVQKIDRKKLKLGIDVSTIEQEL 571
Qy      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
440 --VAVGE-----LTDEDVKMITSLSKDDQIGKEIFASIASPSIYGHEDIKRGPALALFGG 491
Db      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
572 SEQAAGDAEQTRRLTAEEBEKIKRTATRPDLYELLSSSLAPSIVEMDDVKKGILLQLFGG 631
Qy      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
492 EPK--NPGGKHKVRGDNVLLGDGPGTAKSQFLKYIEKVSRAIFTTGCQASAVATYAV 549
Db      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
632 TNKTFQGGNPRYRGDINILLGDPSKSQLRYVHKIAPRGVYTSKGSSAVGLTAYV 691
Qy      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
550 ORHPVSRWETLEAGALVADRGVCLIDFDMKDDORTSTHEAMEQOSISKAGIVTSL 609
Db      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
692 TRDPETHQWLESGALVLDGGICIDEFDKMNSTRSVLHEWEOQTVSIKAGIITL 751
Qy      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
610 QARCTVIAANPIGGRYDPSLTFSENVDLTEPIISRFIDLCVVRDTPVQDEMLARFV 669
Db      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
752 NARTSILASANPIGSRYNPNLPVQNIPLPTLLSRFDLVVLVLDVRVDEQEDRRLAKHLV 811
Qy      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
670 GSHVRHPSNKEEGLANGSAAPAMPNTYGVPELPQEVLLKKYIIYAKERVHKLQMDQ 729
Db      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
812 NMVLEDRPEHAEQ-----ELLPIEFLTAYITYAKTKVHPVLTTPAAG 853
Qy      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
730 DKVAKMYSDLRKES---MATGSIPITVLRHIES 759
Db      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :      :
854 KALSDAYVNRKLGDDTRSSDRITATTRQLES 887

RESULT 11
ID ABJ26305 standard; Protein; 903 AA.
XX
AC ABJ26305;
XX
XX
DT 16-APR-2003 (first entry)
XX
DE Aspergillus fumigatus essential gene protein #963.
XX
XX Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
KW cancer; contamination; biofilm; antibody; immune response.
XX
OS Aspergillus fumigatus.
XX
XX WO200286090-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US13142.
XX
PR 23-APR-2001; 2001US-285697P.
PR 27-APR-2001; 2001US-287066P.
PR 05-JUN-2001; 2001US-295890P.
PR 09-JUL-2001; 2001US-303899P.
PR 31-AUG-2001; 2001US-316362P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX WPI; 2003-093124/08.
XX
XX New purified or isolated nucleic acids of essential genes of
PT Aspergillus fumigatus, useful for treating or preventing infections by
PT A. fumigatus, or for treating a non-infectious disease in a subject
PT e.g. cancer
XX
PS Disclosure; Page -: 175pp; English.
XX
XX The invention relates to novel purified or isolated nucleic acids of

```

```

CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
CC the invention are used to treat or prevent infections by a pathogenic
CC organism such as A. fumigatus, to treat a non-infectious disease in a
CC subject (e.g. cancer), to prevent or contain contamination of an object
CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
CC biofilm comprising A. fumigatus. The polynucleotides are useful for
CC expressing recombinant protein for characterisation, screening or
CC therapeutic use, as markers for host tissues in which the pathogenic
CC organisms invade or reside, for comparing with the DNA sequence of A.
CC fumigatus to identify duplicated genes or paralogues having the same or
CC similar biochemical activity and/or function, for comparing with DNA
CC sequences of other related or distant pathogenic organisms to identify
CC potential orthologous essential or virulence genes, for selecting and
CC making oligomers for attachment to a nucleic acid array for examination
CC of expression patterns, for raising anti-protein antibodies, as an
CC antigen to raise anti-DNA antibodies or to elicit another immune
CC response, and for identifying polynucleotides encoding the other protein
CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case of virulence
CC factors. This sequence represents a protein of one of the essential genes
CC of Aspergillus fumigatus of the invention.
XX
SQ Sequence 903 AA;

```

```

Query Match 19.2%; Score 887; DB 24; Length 903;
Best Local Similarity 33.3%; Pred. No. 2.9e-70;
Matches 230; Conservative 109; Mismatches 252; Indels 100; Gaps 13;

```

```

Qy 221 VNYEDLAAREHVLAYFLPEAPAEQLLQIFDEAALEVLAMYPKYDRIT-NHIVRISHLPL 279
Db 66 VSDIEIRAHNRELADGLTSPEDYSLAFDRALKEVIKTLNRPSTRETTADVNYCAVGA 125
Qy 280 VEEL-----RSLRQLHLNQLRTSGVTSCTGVLPQ-LSMVKYNCKNCFVLGPFCQS--- 331
Db 126 FGEFSCNPRTLGSTHLNRMISLEGIYTKSLVREPKIIQSIVHYNERKORFVARRRDOTMT 185
Qy 332 ----QNEVKPGSCPECQAGPFVNMMEETIYQNYQIRIQESPGKVAARLPRSKDAIL 387
Db 186 TTGTTNNVYP---QEDDEKNPLITEYGYSTYLDHQTISIQEMPERAPAGQLPRSDVIL 242
Qy 388 LADLVDSNAGDEIELTGIYHNNDYDGLSNTANGFPFVATVILANHVAKKDNK----VAVG 443
Db 243 DDDLVDKAKPGDRIQLVGIYRSL--GNRNASSGSSSTERTVVMANNIIQLSSKSGGIAQA 300
Qy 444 ELTDEDVKMITSLSKDDQIGKEIFASTAPSTYGHEDIKRGPALALFGEPEKNPGKHVR 503
Db 301 TTITDTRNINKVAKKNVFEELANSAPSISYGHDIYKTKATLLMLLGGMEKNLDTGTHLR 360
Qy 504 GDINVLCCGDPGTAKSQFLKYIEKVSSRAIFTTGCQASAVATYAVVORHPVSRWETLEAG 563
Db 361 GDINILMVGDPSAKSQLLRFLVNTAPLATATTGSSGVGLTAAYTSDEKTERRELEAG 420
Qy 564 ALVLADRGVCLIDFDMKDDORTSTHEAMEQQOSISKAGIVTSLQARCTVIAAANPIG 623
Db 421 AMVLGDRGVVCIIDEFDKMSDVRVAIHEVMEQQTVTIAKAGIHTSLNARCSVLAAANPIY 480
Qy 624 GRYPDPSLTFSENVDLTEPIISRFIDLCVVRDTPVQDEMLARFVGVSHVHRHPSNKE-- 681
Db 481 QYDPPHDKPHKNIALPDSLSRFDLLFVTDIEDARDVMVSEVHRLMRHVRQGTGEBA 540
Qy 682 -----EEGLANGSAAPAMPNTYGVPELPQE 707
Db 541 PVREDLNQTLGVGLDQNSNQPTVEYKFNVMHLHAGWANSR-----KKGIEILSIP 594
Qy 708 LKYYIIYAKERVHVKPNQMDQDKVAKMYSDLRKESMATG---SIPITVRIHESMSHGGG 764
Db 595 FIKYIQYKSRIPKPVLTGKAADHIVATYSALRNDLSGNGQRTSPITARTLETILRLS- 653
Qy 765 PRAHPSAGLCDDRRRRQHGHPDAGELHRHTVQRHRSNRKTFARYLSFRDNNELLFIL 824

```


PN W0200286122-A2.
XX
PD 31-OCT-2002.
XX
XX 14-MAR-2002; 2002WO-EP03768.
XX
XX 14-MAR-2001; 2001US-275734P.
XX
XX (HYBR-) HYBRIGENICS.
XX
XX Legrain P, Daviet L;
PI WPI; 2003-103412/09.
XX N-PSDB; ACA57364.
XX
XX New complex between two interacting proteins in adipocyte cells, useful
PT for identifying selected interacting domains that modulate protein
PT interactions, or for preventing or treating metabolic disorders such as
PT obesity or diabetes.
XX
XX Claim 6; Page 260-261; 382pp; English.
XX
XX The invention relates to a complex between two interacting proteins in
CC adipocyte cells, given in the specification. The proteins are identified
CC by selecting a bait protein from a known adipocyte marker and then
CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC members of an adipocyte cDNA library. The proteins are designated SID
CC (RTM) (selected interacting domains) proteins. Also included are a
CC polynucleotide encoding a polypeptide in the adipocyte cells, a
CC recombinant host cell expressing at least one of the interacting
CC polypeptides of the complex, selecting a modulating compound in adipocyte
CC cells, a SID (RTM) polypeptide comprising any of the 738 amino acid
CC sequences given in the specification (including its fragment or variant),
CC a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
CC given in the specification (including its fragment or variant), a vector
CC comprising the SID (RTM) polynucleotide, a recombinant host cell
CC comprising the vector, a protein chip comprising the polypeptides and
CC a record comprising all or part of the data, listed in the specification.
CC The complex, polypeptides, polynucleotides and compounds are
CC useful for preventing or treating metabolic disorders such as obesity
CC or diabetes. The polynucleotides are useful as probes or primers. The
CC complex is particularly useful for identifying selected interacting
CC domains (SID (RTM)) for screening drugs that modulate the protein
CC interaction, thus exhibiting the therapeutic effect. The present
CC sequence represents a SID (prey) protein of the invention.
XX
SQ Sequence 167 AA;
Query Match 18.9%; Score 875; DB 24; Length 167;
Best Local Similarity 100.0%; Pred. No. 2.2e-70;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 205 KERISDMCKNRESLVNVEDLAAREHVLAYFLPEAPAEQLQIFDEAALEVVLAMYPKYD 264
DB 1 KERISDMCKNRESLVNVEDLAAREHVLAYFLPEAPAEQLQIFDEAALEVVLAMYPKYD 60
QY 265 RITNHHIVRISHLPLVEELRSRLQLHNLQIRTSQGVVTSCTGVLPLQSLMVKYKNCNFV 324
DB 61 RITNHHIVRISHLPLVEELRSRLQLHNLQIRTSQGVVTSCTGVLPLQSLMVKYKNCNFV 120
QY 325 LGPFCOSQNVKPGSCPCQAGPFEVNMETIYQNYQIRIQESP 371
DB 121 LGPFCOSQNVKPGSCPCQAGPFEVNMETIYQNYQIRIQESP 167
RESULT 14
AAB95203
ID AAB95203 standard; Protein; 863 AA.
XX
AC AAB95203;
XX
DT 26-JUN-2001 (first entry)
XX

DE Human protein sequence SEQ ID NO:17301.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs.
XX
XX Claim 8; SEQ ID 17301; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 863 AA;
Query Match 18.8%; Score 869; DB 22; Length 863;
Best Local Similarity 30.3%; Pred. No. 1.1e-68;
Matches 245; Conservative 128; Mismatches 282; Indels 154; Gaps 23;
QY 3 SSPAQRNRGND-----PLTSSPG---RSSRRTDALTSPPGRDLPPFEDESEGLGTE 51
DB 31 SSPSQRNRGDSSTGELQPMPTSPGVDLQSPAQVLFSSP----- 72
QY 52 GPLEEEEDGEELIGDNERDYRAIPELDVAEGLALDDDEDVELT-----ASRREA--D 105
DB 73 -----PQM-----HSSAIPLDFDVSSPLTYGTSPVEGTPRS 105
QY 106 GPCGTVTGSWPGLCAGVCGSCMTAMRRTRRSALPASAGAGTGEDGEDEQMIENLED 165
DB 106 GVRGTPVRQRPDLGSAQKGLQVDLQ-----SDGNAAREDIVASEQSL----- 146

```
QY 166 LKGHSVREW---VSMAGPRLRIHHRFKNFLRTHVD---SHGHNV-----FKERISDM 211
Db 147 --GOKLVITGTDVNAACK----ENQFQRLQRTIDPLAKEENGVNIDITEPLYMQRLGEI 200
QY 212 CKNRESLVNYEDLAAREHVLAFLPEAPAEQLLOIFDEAALEVLVWAMPKYDRINH-I 270
Db 201 NVTGEPLVNCHEIKSFQDKNLYRQLISYPOEVIPTFDMAVNEIFFDRYP--DSILEHQI 258
QY 271 HVRISHLPLVEELRSRLQLHLNOLIRTSGVVTSCTGVLPOLSMVKYCNCKNCFVLPFCQ 330
Db 259 QVRPFNALKTKNRNLNPEDIDOLITISGVNIRTSQILPEMQEAFQCCQCAHTT-RVEM 317
QY 331 SQNOEYKPGSCPCQAGPEVNMETIYQYQIRIQESPGKVAARRLPRSKDAILLA- 389
Db 318 DRGRIAEPSVCGCHTTHSMALIHNRSLFSDQMIKLQESPEMPAGQTPT--VILFAH 375
QY 390 -DIVDSNAGDETELGIYH-----NNYDGLSNTANGFPVFATVILANHVAKDKNVAV 442
Db 376 NDLDVKVQGDVNVNVTGIYRAVPIRNP-----RVSNVKSYYKTHIDVTHYKTDKRLH 430
QY 443 GELTDED-----VKMITSLSKDQOIGERKIFASIPSIYGHEDIKRGPALALFGGE 492
Db 431 G--LDEEAQKLFSEKRVELLKELSRKPDYIERLASALAPSIYEHEDIKGIILLQFEGT 488
QY 493 PK--NPGGKHVKGDNVILCGDPAKQFLYIEKVSSRAITFTGQGASAVATYVQ 550
Db 489 RKDFSHTGRGKFAEINILCGDPTGSKSOLLOYVNLVPRGQYTSKGSSAVGLTAYVM 548
QY 551 RHPVSREWTLEAGALVADRGVCLIDFEFKMNDORTSIHEAMEQOSISISKAGIYTSQ 610
Db 549 KDEPTELVLQTLGALVSDNGICIDEFDRNNESTRSVLHEVMEQOTLSAKAGIICQLN 608
QY 611 ARCTIAAANPIGGRYDPSLTFSENVDLTPETIISRFEDILCVWRDTPVPODEMALARFVG 670
Db 609 ARTSVLAANPIESQNNPKTTIENIQLPHTLSRFDLIFMLDPQDEAVDRRLAHLVA 668
QY 671 SHVRHPSNKEEGLANGSAEPAMPNTYGVPELPQEVLVKKYIYAKERVHPKLNQMDQD 730
Db 669 LTYQSEEQAE-----ELLDMAVLKDYIAYARHSTIMPRLSSEASQ 708
QY 731 KVAKMYSDLRKESMATGSPITVVRHIESM 759
Db 709 ALIEAVDMRKIGSSRGMSAYPRQLES 737
RESULT 15
ID ABJ25900
XX ABJ25900 standard; Protein: 874 AA.
XX AC ABJ25900;
XX XX 16-APR-2003 (first entry)
XX XX Aspergillus fumigatus essential gene protein #558.
XX DE Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
XX KW cancer; contamination; biofilm; antibody; immune response.
XX XX Aspergillus fumigatus.
XX XX W0200286090-A2.
XX PN 31-OCT-2002.
XX PD 23-APR-2002; 2002WO-US13142.
XX PF 23-APR-2001; 2001US-285697P.
XX PR 27-APR-2001; 2001US-287066P.
XX PR 05-JUN-2001; 2001US-295890P.
XX PR 09-JUL-2001; 2001US-303899P.
XX PR 31-AUG-2001; 2001US-316362P.
XX XX (ELIT-) ELITRA PHARM INC.
PA
```

```
XX Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
PI WPI; 2003-093124/08.
XX New purified or isolated nucleic acids of essential genes of
PT Aspergillus fumigatus, useful for treating or preventing infections by
PT A. fumigatus, or for treating a non-infectious disease in a subject
PT e.g. cancer
XX Disclosure; Page -: 175pp; English.
XX The invention relates to novel purified or isolated nucleic acids of
CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
CC the invention are used to treat or prevent infections by a pathogenic
CC organism such as A. fumigatus, to treat a non-infectious disease in a
CC subject (e.g. cancer), to prevent or contain contamination of an object
CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
CC biofilm comprising A. fumigatus. The polynucleotides are useful for
CC expressing recombinant protein for characterisation, screening or
CC therapeutic use, as markers for host tissues in which the pathogenic
CC organisms invade or reside, for comparing with the DNA sequence of A.
CC fumigatus to identify duplicated genes or paralogues having the same or
CC similar biochemical activity and/or function, for comparing with DNA
CC sequences of other related or distant pathogenic organisms to identify
CC potential orthologous essential or virulence genes, for selecting and
CC making oligomers for attachment to a nucleic acid array for examination
CC of expression patterns, for raising anti-protein antibodies, as an
CC antigen to raise anti-DNA antibodies or to elicit another immune
CC response, and for identifying polynucleotides encoding the other protein
CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case of virulence
CC factors. This sequence represents a protein of one of the essential genes
CC of Aspergillus fumigatus of the invention.
XX SQ Sequence 874 AA;
Query Match 18.6%; Score 860; DB 24; Length 874;
Best Local Similarity 29.6%; Pred. No. 7.5e-68;
Matches 251; Conservative 127; Mismatches 269; Indels 200; Gaps 25;
QY 2 ASSPA-----QRRRGNDPLTSS-----PGRS-SRRDT-----ALTSPPCRDLPPF 40
Db 3 SSSPTRAQNQRSSRDIPSSSGLFVSSRPSIESNRAVRSRDLHSGGFLSSNRRRVF 62
QY 41 EDESEGLGTGTEGLEEEEDGEELIGDGMERDYRAIPELDAYEAEGALDDEDEVELTASR 100
Db 63 VD-ANGMPATDGD-----PRSDA-TFSNIHPDTSEAEALGSS 98
QY 101 REAADPGCTVTCGFWPLGACAGVSCWTAMR-----RTRSPALPASAACTGDEGEDE 154
Db 99 TRVITGNTSIQDS-----MSAFKNFLYNFQTKYRLMAEGA----TED--ETR 140
QY 155 QMIESIENLEDLKGHSVREWVSMAGPRLRIHHRFKNFLRTHVDVSHGHNVFKERISDMCKE 214
Db 141 IMGDSAE-----REYISM-----LSTMRLQ 161
QY 215 NRESLVVNYEDLAAREHVLAFLPEAPAEQLLOIFDEAALEVLVWAMPKY-----YD 264
Db 162 GVTSLNDAKNLAKAYPSTLKLWHQLHAYPOEIIPLMDQTVKDVWVVELAIKEMERLRAQ 221
QY 265 RITNH-----IHRVISHLPLVEELRSRLQLHLNOLIRTSGVVTSCTGVLPOLSM 313
Db 222 RNQNHRLGLSSVETKAFKVLPLFGLDSTVNMRLDPADMDKLVSKIGLIVIRTTPIPDKE 281
QY 314 KYNCNCKNFVLGPFQCSQNOEVKPGSCPE--COSAGPFVNMETIYQYQIRIQES 371
Db 282 AFFRCQVCHGV-QVDIDRGKIAETPCRPVCKERNMQLIHNRVCFADKQVLIKQETP 340
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2003, 13:10:40 ; Search time 26 Seconds
(without alignments)
3299.326 Million cell updates/sec

Title: CAA47749
Perfect score: 4630
Sequence: 1 MASSPAQRRRGNDPLTSSPG.....LFRMNFHDLKRKMLQOF 892
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4630	100.0	892	1 S42228	replication licens
2	3941	85.1	904	1 T10067	replication licens
3	3539.5	76.4	886	1 JC5085	replication licens
4	2254	48.7	881	2 T26498	hypothetical prote
5	1951.5	42.1	830	1 B48723	replication licens
6	1945.5	42.0	936	2 E96508	hypothetical prote
7	1812	39.1	868	1 S45757	replication licens
8	913	19.7	686	2 H90226	minichromosome mai
9	909.5	19.6	699	2 B72775	probable DNA repli
10	877	18.9	863	1 S65954	replication licens
11	874.5	18.9	863	1 S64720	replication licens
12	873	18.9	879	2 T41059	replication licens
13	858	18.5	720	2 C84540	probable CDC21 pro
14	856	18.5	862	1 S56766	replication licens
15	855	18.5	776	2 T52118	probable replicati
16	848.5	18.3	933	1 S56050	replication licens
17	843.5	18.2	858	2 T47223	replication licens
18	843	18.2	909	1 S26640	replication licens
19	825.5	17.8	735	1 PC4225	replication licens
20	823.5	17.8	666	1 D69103	DNA helicase (EC 3
21	819.5	17.7	866	1 S59872	replication licens
22	807.5	17.4	796	2 T47222	replication licens
23	802	17.3	971	1 A36376	replication licens
24	796.5	17.2	727	2 G84487	probable DNA repli
25	793.5	17.1	1017	1 S64219	replication licens
26	789	17.0	775	1 A39631	replication licens
27	787.5	17.0	734	1 I38080	replication licens
28	786.5	17.0	892	2 T50339	mis5 protein fimo
29	784	16.9	586	1 E69314	replication licens

ALIGNMENTS

RESULT 1

S42228

replication licensing factor MCM2 - human

N;Alternate names: CDCL1; minichromosome maintenance protein MCM2 homolog; protein BM

C;Species: Homo sapiens (man)

C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 20-Aug-1999

C;Accession: S42228

R;Odorov, I.T.; Pepperkok, R.; Philippova, R.N.; Kearsey, S.E.; Ansorge, W.; Werner, J. Cell Sci. 107, 253-265, 1994

A;Title: A human nuclear protein with sequence homology to a family of early S phase

A;Reference number: S42228; MUID:94230605; PMID:8175912

A;Accession: S42228

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-892 <TOD>

A;Cross-references: EMBL:X67334; NID:9468703; PIDN:CAA47749.1; PID:9468704

C;Comment: The complex of six MCM proteins is one of several proteins that must be bo phosphorylated and dissociate from the chromatin.

C;Genetics:

A;Gene: GDB:MCM2; CDCL1; D3S3194; cdc19

A;Cross-references: GDB:224876; OMIM:116945

A;Map position: 3q21-3q21

C;Complex: The predominant form is a heterohexamer of MCM2, MCM3, MCM4, MCM5, MCM6, a

C;Function:

A;Description: part of the replication licensing system that permits DNA replication

C;Superfamily: replication licensing factor MCM2; MCM homology

C;Keywords: cell cycle control; DNA replication initiation; heterohexamer; nucleus; p

F;283-786/Domain: MCM homology <MCM>

F;318-344/Region: zinc finger CCCC motif

Query Match 100.0%; Score 4630; DB 1; Length 892;

Best Local Similarity 100.0%; Pred. No. 1.3e-285;

Matches 892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Db 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Qy 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Db 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Qy 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Db 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Qy 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Db 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Qy 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Db 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Qy 100.0%; Score 4630; DB 1; Length 892;

Best Local Similarity 100.0%; Pred. No. 1.3e-285;

Matches 892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Db 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Qy 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Db 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Qy 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Db 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Qy 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Db 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Qy 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Db 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Qy 100.0%; Score 4630; DB 1; Length 892;

Best Local Similarity 100.0%; Pred. No. 1.3e-285;

Matches 892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Db 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Qy 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Db 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Qy 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Db 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Qy 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Db 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Qy 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Db 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Qy 100.0%; Score 4630; DB 1; Length 892;

Best Local Similarity 100.0%; Pred. No. 1.3e-285;

Matches 892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Db 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Qy 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Db 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Qy 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Db 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Qy 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Db 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Qy 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Db 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Qy 100.0%; Score 4630; DB 1; Length 892;

Best Local Similarity 100.0%; Pred. No. 1.3e-285;

Matches 892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Db 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Qy 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Db 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Qy 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Db 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Qy 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Db 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Qy 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Db 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Qy 100.0%; Score 4630; DB 1; Length 892;

Best Local Similarity 100.0%; Pred. No. 1.3e-285;

Matches 892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Db 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Qy 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Db 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Qy 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Db 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Qy 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Db 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Qy 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Db 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Qy 100.0%; Score 4630; DB 1; Length 892;

Best Local Similarity 100.0%; Pred. No. 1.3e-285;

Matches 892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Db 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Qy 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Db 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Qy 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Db 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Qy 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Db 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Qy 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Db 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Qy 100.0%; Score 4630; DB 1; Length 892;

Best Local Similarity 100.0%; Pred. No. 1.3e-285;

Matches 892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Db 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Qy 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Db 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Qy 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Db 121 CAVGSCMTAMRRTRSALPASAAAGAGTDEDEEQMIESTENLEDLKGHSVREWSWAGP 180

Qy 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Db 181 RLEIHRFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA 240

Qy 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Db 241 PAELLQITFDEAALVEVLAMYPKYDRITNHHVRIHSLPLVEELRSRLQLHNLIRTSVG 300

Qy 100.0%; Score 4630; DB 1; Length 892;

Best Local Similarity 100.0%; Pred. No. 1.3e-285;

Matches 892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Db 1 MASSPAQRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

Qy 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREAADGPGCTVTSWPGIGA 120

Db 61 EELIGDGMERYRAIPELDADYEAGLADDEDEVEELTASRREA

Db 363 AGRIPRSKDVILLGDLCDSCKPGDEIEVTGYTNNFDSGLNYKQGFVPVNTLIHANHTN 422
Qy 436 KDNKVAVGELTDEVDKMTSLSKDQOIGEKIFASTAPSIYGHEDIKRGPALALFGGEPKN 495
Db 423 KD-KWASDQLDEDEIKAIRESQDPSNTSQRVFSIASIYGHDDVKRAIALALFRGEAKN 481
Qy 496 PGGKHKVRGDIINVLLCGDPGCTAKSQFLKYIEKVSRAIFTTGGQASAVAVTAYVQRHPVS 555
Db 482 PGAKHRLRGDINVLLCGDPGCTAKSQFLRYAAHIAPRSVLTGQASAVGLTAYVQRHPVT 541
Qy 556 RENTLEAGALVLDARGVCLIDEFKMNDORTSHEAMEQOSISISKAGIVTSLOARCTV 615
Db 542 RENTLEAGAMVLDARGVCLIDEFKMSQDRTSHEAMEQOSISISKAGIVTSLFARCTV 601
Qy 616 IAAANPIGGRYDPSLTSENVDLTPETISRDILCVVDRVDPQVDEMALARFVGVSHVRH 675
Db 602 IAAANPIGGRYNPRTFAENVLDLTPILSRFDVLCVINDSVDSVEDERLAKFVGVGNRHTH 661
Qy 676 HPSNKE--EEGLANGSAAPAMPNTYGVPELPQEVLYKIIYAKERVHPKLNQMDQKVA 733
Db 662 HPDAKKIVKEG--DELEEDQMDERTGVRLIPQDLRLKRIIYAREKCHPTLPEQHSEKFS 718
Qy 734 KMSDLRKESMATSGIPITVYRHIESMSHGGGPRAHPSNGLCDRRRRQHGHPRDAGE---- 789
Db 719 NIFAQMRKESMATSGVAITVYRHVESMIRLS--EAAKHLH-----RSYVNDEDCAAAIRV 771
Qy 790 -LHRHTEVQR---HRSMRKTFRARYLSFRDNNEILLFLKOLVAEOVY--ORNRFGAQO 843
Db 772 MLESFVNTQKASIMRMKKFTSRHLTENRSANELLLFLKOLIRQOHHYATARAAGTIL 831
Qy 844 DTIEVPEKDLVDKARQINHLNSAFYDSELFPMKFKSHDLKRRKMILQO 891
Db 832 QSVTPESEFTEKAQQLRIENVKPFYTFEISFASNFLYDPSKKTIVQE 879

RESULT 5
B48723
Replication licensing factor MCM2 homolog ndal - fission yeast (Schizosaccharomyces pombe)
A:Alternate names: cell division cycle protein cdc19
C:Species: Schizosaccharomyces pombe
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 18-Jul-2001
C:Accession: B48723; T40418; T52476
R:Miyaake, S.; Okishio, N.; Samejima, I.; Hiraoka, Y.; Toda, T.; Saitoh, I.; Yanagida, M. Mol. Biol. Cell 4, 1003-1015, 1993
A:Title: Fission yeast genes ndal(+) and nda4(+), mutations of which lead to S-phase block
A:Reference number: A48723; MUID:94129084; PMID:8298187
A:Accession: B48723
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-830 <MIY>
A:Cross-references: GB:S68468; NID:g545212; PIDN:AAC60569.1; PID:g545213
R:McDougall, R.C.; Rajandream, M.A.; Bartell, B.G.; Brown, S.; Harris, D. submitted to the EMBL Data Library, October 1999
A:Reference number: 221927
A:Accession: T40418
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-830 <MCD>
A:Cross-references: EMBL:AL121863; PIDN:CAB58403.1; GSPDB:GN000067; SPDB:SPBC4.04c
A:Experimental source: strain 972h; cosmid c4
R:Forstburg, S.L.; Nurse, P.
J. Cell Sci. 107, 2779-2788, 1994
A:Title: The fission yeast cdc19+ gene encodes a member of the MCM family of replication
A:Reference number: 226087; MUID:95181580; PMID:7876346
A:Accession: T52476
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-830 <FOR>
A:Cross-references: EMBL:U08048; PIDN:AAC48930.1
C:Comment: The complex of six MCM proteins is one of several proteins that must be bound phosphorylated and dissociate from the chromatin.
C:Genetics:
A:Gene: ndal+

A:Map position: 2
A:Introns: 78/2
A:Note: cdc19+
C:Complex: The predominant form is a heterohexamer of MCM2, MCM3, MCM4, MCM5, MCM6, a C:Function:
A:Description: part of the replication licensing system that permits DNA replication
litedated, MUID:95181580)
C:Superfamily: replication licensing factor MCM2; MCM homology
C:Keywords: cell cycle control; DNA replication initiation; heterohexamer; nucleus; p
F:298-804/Domain: MCM homology <MCM>
F:334-360/Region: zinc finger CCCC motif

Query Match 42.1%; Score 1951.5; DB 1; Length 830;
Best Local Similarity 48.2%; Pred. No. 9.7e-116;
Matches 418; Conservative 116; Mismatches 238; Indels 95; Gaps 12;

Qy 1 MASSPAQRRCNDPLTSSPGRSSRRTDALTSPPGRDLPPPEDESEGLLTGEGPLE----- 55
Db 1 MDSEFKRGRDSESLPFSESENSSLGATPLSLPPSPPPPEFSDEAAEAL-VEEDIEDLDGE 59
Qy 56 ---EEDEGELIGDMERDYRAIPELDAYEAGLALDDEVEELTASRRRAADGPGCTV 111
Db 60 ALDVEDEGEDLFGEGMERDYQONLEDRYDIELD-DDNLEELDICARRAVD----- 112
Qy 112 TGSWPGLCACAVGSCMTAMRRTSRALPASAA-----SGAGT----- 147
Db 113 -----ARLRRDIELDAAAGRTKPAAFLOQEDDDLDLSNLCTGFTRRHR 156
Qy 148 -----EDGEEDROMIESIENEDLKGHSVREWVSMAGPRLIHHKPKNLTHTVDSHG 200
Db 157 IYDEYSPNVGALDESSELPLESIADVADSAEAWTLDPVVRTTAREFKNLFLEYTDENG 216
Qy 201 HNVKERISDMCKENRESLVVYEDLAAREHVLAYFLPEAPAEALLOQLIDEAELVVLAMY 260
Db 217 TSVYGNRIITLGEVNAESLVVYAHLGESKPIAYFLANAPATFRIDRVALATLHY 276
Qy 261 KYDRITNHHVRISHLPVBEELSLRLQLHNLQIRTSVGVVTSCTGVLPLQSLVMYKNCNK 320
Db 277 PDYERIHSDIHVRTNLPTCTFLDLRQSHLNCULVRVSGVVVTRTGLPQLKYIRFTCTK 336
Qy 321 CNFVLGPFQCSQNOEVRPGSCPEQCSAGPFEVNNMEETIYQYORIRIOESGKVAARLPL 380
Db 337 CGATLGPFFQDSSVEVAKISFCNCHSSRGPPFINSERTVYNNYQRITQESGTPVSGKLP 396
Qy 381 RSKDAILLADLVDSNAGDETELFGIYHNNYDGSINTANGPFPFATVTLAHNAK----- 435
Db 397 RHREVILLADLVAKPGEEIDVTGIYRNNEFASLNTKNGFPFATIEANHISQLDGS 456
Qy 436 -KONKAVGELTDEEDVMTSLSKDQOIGEKIFASIPSIYGHEDIKRGPALALFGGEPK 494
Db 457 NTDDDFSLSLRUTDDEEREIRALAKSPDIHNRRIIASMAPSIYGHRSIKTAIAAALFGGVPK 516
Qy 495 NPGCKHKVRGDIINVLLCGDPGCTAKSQFLKYIEKVSRAIFTTGGQASAVAVTAYVQRHPV 554
Db 517 NINGKHIRGDIINVLLGDPGCTAKSQFLKYIEKVAHRAVAVFATGOGASAVGLTASVRKDP 576
Qy 555 SREWTLEAGALVLDARGVCLIDEFKMNDORTSHEAMEQOSISISKAGIVTSLOARCT 614
Db 577 TNEWTLEGGALVLDARGVCLIDEFKMNDORTSHEAMEQOSISISKAGIVTLOARCT 636
Qy 615 VIAANPIGGRYDPSLTSENVDLTPETISRDILCVVDRVDPQVDEMALARFVGVSHVR 674
Db 637 IIAAANPIGGRYNTTIPFNQNVLETILSRFDILQVVKDVPNPEIDQLANFVNVSSHIR 696
Qy 675 HPSNKEEGLANGSAAPAM-----PNTYGVG--PLPQEVLYKIIYAKERVHPKLNQ 726
Db 697 SHP-----AFDPNMDVLKVKVPTETGIDAKPIQDLRLKRIIYAREKCHPTLPEQH 744
Qy 727 MDQKVAKMYSDLRKESMATSGIPITVYRHIESMSHGGGPRAHPSNGLCDRRRRQHGHPRD 786
Db 745 MDEEKISRLYSMDMRRESLATGSPITVYRHLESARLS--EAFKMWLSEFVRPSSHIDKAI 802
Qy 787 AGEUHRHTEVQR---HRSMRKTFRARYL 810

A:Cross-references: GB:X53539
C:Comment: The complex of six MCM proteins is one of several proteins that must be bound to DNA to initiate DNA replication. The complex is phosphorylated and dissociate from the chromatin.
C:Genetics:
A:Gene: SGD:MCM2; MIPS:YBL023c
A:Cross-references: SGD:S0000119; MIPS:YBL023c
A:Map position: 2L
C:Function:
A:Description: MCM2 is a component of the replication licensing factor that permits DNA replication.
C:Superfamily: Replication licensing factor MCM2; MCM homology
C:Keywords: cell cycle control; DNA replication initiation; heterohexamers; nucleus; phosphatase; zinc finger
F:306-840/Domain: MCM homology <MCM>
F:341-367/Region: zinc finger CCCC motif

Query Match 39.1%; Score 1812; DB 1; Length 868;
Best Local Similarity 46.3%; Pred. No. 7.3e-107;
Matches 384; Conservative 112; Mismatches 222; Indels 112; Gaps 9;
QY 3 SSPAQRGRNDPLTSSPGRRSRRTDALTSFGRDLPPFEDESEGLLTGTEPLEEEDGEE 62
DB 23 SSPQHFGRGMPVSSP-----IGSPDMINPEGDNDVDDVPDIDEVEEQMNEVD 72
QY 63 LIGDQWERYRAIPELDAYEAGLALDDDEYBELTASRREAGDPCGTVTGWSWGLGACA 122
DB 73 LMDNNYEDYAAADHNRDRYDPD--QVDDREQELSLSERRID----- 113
QY 123 VGSCTAMRRRTSALPASAGSAGTDEGDEEQ-----MIESTEN 162
DB 114 -----AQINERDRLRNVA-----YIDDEDEQEGAAOLDDEMGLPVQRRRRQVEDLEN 163
QY 163 LED-----LKGHSVREWYSMAGPRLIHRHKNFLRTHVDSHG 200
DB 164 SDDLLSDMDIDPLEELTLESLSNVKANSYSEWITQPNVSKRTIARELKSFLLEVDTGT 223
QY 201 HNVFERSIDMCKENRESLVNVEDLAAREHVLAYFLPEAPAEALLOIFDEAALEVVLAMY 260
DB 224 RSVYGARITLGE MNSESLVYRHLAESKAILALFLAKCPPEMLKIFDLVAMEATELHY 283
QY 261 PKYDRITHIHRVSHLPLVBEELSLRQLHNLQIRTSQVVTCTGVLPSQSMKYKNC 320
DB 284 PDYARHSEIHVRISDFTIYSLRELRESNLSLVRTGVVTRRTGVFPQLKYKFNCKL 343
QY 321 CNFVLGPGCSQNGQVKGPCQSCQAGPFFVNMETIYQNYQRIQIESPGKVAARLLP 380
DB 344 CGSILGPFQDSNEBIRISFCTNCKSGPFRVNGEKTYYRNYQVTLQEAQCTVPPGRLP 403
QY 381 RSKDAILLADLVDSNAGDELTGTYHNNVDGSLTANGPFFVATVILANHVAKKNKV 440
DB 404 RHREVTLLADLVDSKPGEEVEVTGIYKNNYDGNLNKNGPFPVATIANSIKRRGNT 463
QY 441 A-----VGELTDVDMKITSLSKQOQIGEIFASTAPSIYGHEDIKRGPPALALFGCE 492
DB 464 ANEGEGLDVFSWTEEEERERKTSRDRGIIDKIISNAPSIIYGHRIKTAACSLFGV 523
QY 493 PKNPGKHVIRGIDINVLGCGDPGTAQSOFLYKIEKVSSRAIFTTGGQASAVATYAVQRH 552
DB 524 PKNVNGKHSIRGIDINVLGCGDPGTAQSOFLYKIEKVSSRAIFTTGGQASAVATYAVQRH 583
QY 553 PVSREWTLEAGALVADRGVCLIDFDFKNDODRTSIHEAMEQOSISISKAGIYVLSOAR 612
DB 584 PITKEWTLLEGALVADRGVCLIDFDFKNDODRTSIHEAMEQOSISISKAGIYVLSOAR 643
QY 613 CTVTAANPIGGRYDPSLTFSNVDLTPITSRFDLCVWRDTPVQDEMELARFVGS 672
DB 644 CSIIAANPNNGRYNSTPLAQNVSLTEPILSRFDLCVWRDTPVQDEMELARFVGS 703
QY 673 VRHHSNKE----EGLANGSAA-----EPAMPNTYGVPELPQBV 709
DB 704 VRSHPENDEDEGEELKNGSESATQEGDEINEQLNQRRLQORRKEEISIPQBEL 763
QY 710 KKYIYAKERVHVKLNQMDQKVAKMYSDLRKESMATGSIPTVRRHISM 759
DB 764 MKYIYAKRTKYPKLQMDQKVSKVYADLRRESISTGTSFPTVRRHLESI 813

RESULT 8

H90226
minichromosome maintenance protein MCM [Imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: H90226
Jong, I.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Ch
She, Q.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: H90226
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-866 <KUR>
A:Cross-references: GB:AE006641; NID:gl3813947; PIDN:AAK41071.1; GSPDB:GN00155
C:Genetics:
A:Gene: SSO0774

Query Match 19.7%; Score 913; DB 2; Length 686;
Best Local Similarity 36.2%; Pred. No. 5.6e-50;
Matches 216; Conservative 115; Mismatches 215; Indels 50; Gaps 11;
QY 180 PRLEIHRH--EKNFLRTHVDSHGHNVPKERSIDMCKENRESLVNVEDLAAREHVLAYFL 237
DB 4 PSKQIDYRDVIEFLTFEGNNQNNKYIERINELVAYRKSLIIEFSVLSFNENLAYEI 63
QY 238 PEAPAEALLOIFDEAALEVVLAMYPKYDRITHIHRVSHLPLVBEELSLRQLHNLQIR 297
DB 64 INNTKIILPELEGALYDHILQDPTYQRIEDIEKVHVRIGVPRVIELRIRSTIDIGKITI 123
QY 298 SGVVTSCITGVLPQLSMWKYNKNCNVLGPGCS-----QNEV-----KGSCEPCQSA 347
DB 124 DGIIVKVPVVKRIYKATYK-----HIHPDQMEFEWPEDEEMPEVLEMTPTICPKCKP 177
QY 348 GPFEVNMETIYQNYQRIQIESPGKVAARLLPRSKDAILLADLVDSNAGDEIETGY 407
DB 178 GQRLIPEKTKLIDWQKAVIQERPEVPSGQLPROLEILDEDDLVDSARPGRVKTGIL 237
QY 408 HNNVDGSLTANG--FPVFATVILANHVAKKNKAVANGELTDEEDVKMTISLSKQOIGEK 465
DB 238 DIKQDSPVKGRSRAVFDIYMKVSVSEVSQKVLDEVIIE--EDEKKIKDLAKDPWIRD 294
QY 466 IFASTAPSIYGHEDIKRGPPALALFGGEPKNGKHVRGDIINVLGCGDPGTAQSOFLY 525
DB 295 IISSTAPSIYGHWEUKALALFGGVPKVLEDT--RIKGDHILIIIGDPGTAQSOMLQFI 353
QY 526 EKVSRAIFTTGGQASAVATYAVQRHVPVSREWTLLEAGALVADRGVCLIDFDFKNDQD 585
DB 354 SRVAPRAVYTTGKSTAGLTAAVVREKGTGEYVLEAGALVADRGVCLIDFDFKNDQD 413
QY 586 RTSIHEAMEQOSISISKAGIYVLSOARCTVTAANPIGGRYDPSLTFSNVDLTPITSR 645
DB 414 RVATHEAMEQOVTSTAKAGIVAKLNARAATAAGNPKFGRVYSRPSVDNINLPPTILSR 473
QY 646 FDLICVWRDTPVQDEMELARFVGSVHRRHPSNKEEGLANGSAAEPAMPNTYGVPELP 705
DB 474 FDLIFILQDQGE--ODRELANYILDVH-----SGKSTKNI-----ID 509
QY 706 QEVLLKYYIYAKERVHVKLNQMDQKVAKMYSDLRKESMATGSIPTVRRHISM 759
DB 510 IDTLRKYYIYAKERVHVKLNQMDQKVAKMYSDLRKESMATGSIPTVRRHISM 565

RESULT 9

B72775
probable DNA replication licensing factor mcm APE0188 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: B72775

QY	331	SNQEVKPCSCPCQAGPFEVMEETIYONYQIRIQIESPGKVAARRUPRSKDAILLA	389
Db	318	DRGRIAEPSVCGRCHTTHSMALIHNRSLSFDKQMIKLQESPEDMPAGQTPHT--VILFAH	375
QY	390	DLVDCSNAGDETELTGIYH-----NNDGSLNTANGPPVFATVILANHVAKDNKVAV	442
Db	376	NDLVDKVPQGDVNNVTGIVRAVPIRNP-----RVSNVKSVMYKTHIDVIRYKRTDAKRLH	430
QY	443	GELTDED-----VKMTLSLKDQOIGEKIFASIAPSIYGHEDIKRGPALALFGGE	492
Db	431	G--LDGEAQKLFSEKRVKVELLKSRRPDIYEKLASALAPSIYEHEDIKGIILLQFGGT	488
QY	493	PK--NPGCKHKVRGDNVLLCGDPGTAQSOFLKYIEKVSSRAIFFTTGGASAVAVTAVQ	550

Db 489 RKDFSHTGRGKFAEINILLCGDPGTSKSQLQVYVNLVPRGQVTSKGSSAVGLTAYVM 548
 Qy 551 RHPVSRWTLGAGALVLDRCVGLIDFEDKNDODRTSIHEAMBOQOSISIKAGIVTSIQ 610
 Db 549 KDPETRLQVLOTGALVLDSDNGICCIDFEDKNESTRSVLHEVMEQQVLSIAKAGIICQLN 608
 Qy 611 ARCTVIAAANPIGGYRPSLTFSENVDLTPETIISRFIDILCVVRDTPVQDMLARFVVG 670
 Db 609 ARTSVLAAANPIESQWNPCKTTIENIQLPHTLLSRFDLFLMDPQDEAYDRRLAHLVA 668
 Qy 671 SHVRHSHNSKEEGLANGSAAPAMPNTYGVPELPQEVLLKYYIYAKERVHPKLNQMDQD 730
 Db 669 LYYQSEEQAE-----ELDDMAVLKDYIAAHSTIMPRLSEASQ 708
 Qy 731 KVAKWYSLDRKESMATGIPITVRHIESM 759
 Db 709 ALIEAYVDMRKIGSRGMVSAYPROLES 737

RESULT 11
 S64720
 N:Alternation: licensing factor MCM4 - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000
 C:Accession: S64720; S26643; S25529
 R:Coou, M.; Kearsey, S.E.; Mechali, M.
 EMBL J. 15, 1085-1097, 1996
 A:Title: Chromatin binding, nuclear localization and phosphorylation of Xenopus cdc21 and
 A:Reference number: S64720; MUID:96183193; PMID:8605878
 A:Accession: S64720
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-863 <COU>
 A:Cross-references: EMBL:U29178; NID:g1002597; PIDN:AA01680.1; PID:g1002598
 R:Coxon, A.; Maundrell, K.; Kearsey, S.E.
 Nucleic Acids Res. 20, 5571-5577, 1992
 A:Title: Fission yeast cdc21(+) belongs to a family of proteins involved in an early step
 A:Reference number: S26640; MUID:93087163; PMID:1454522
 A:Accession: S26643
 A:Molecule type: DNA
 A:Residues: 513-523, 'Y', 525-538, 'G', 540-553, 'G', 555-588 <COX>
 A:Cross-references: EMBL:Z15033; NID:g64612; PIDN:CAA78751.1; PID:g64613
 C:Comment: The complex of six MCM proteins is one of several proteins that must be bound
 phosphorylated and dissociate from the chromatin.
 C:Genetics:
 A:Gene: cdc21
 C:Complex: The predominant form is a heterohexamer of MCM2, MCM3, MCM4, MCM5, MCM6, and
 C:Function:
 A:Description: part of the replication licensing system that permits DNA replication to
 C:Superfamily: replication licensing factor MCM4; MCM homology
 C:Keywords: cell cycle control; DNA replication initiation; heterohexamer; nucleus; phos
 P:271-764/Domain: MCM homology <MCM>

Query Match 18.9%; Score 874.5; DB 1; Length 863;
 Best Local Similarity 28.1%; Pred. No. 2.1e-47;
 Matches 264; Conservative 145; Mismatches 358; Indels 173; Gaps 25;

Qy 3 SSPAQR-----RGNDPLT-----SSPGRSRRDALT-----SPGRDLPPEDE 43
 Db 6 STPSRRNRKRGSNPTPHGEEVQSPQRRRTSTGELLPMPTSPSGDV----- 59
 Qy 44 SEGLLGTEGPEEEDGEEELGDMRDYRAIPELDVAEAGLADDDVEELTASRREA 103
 Db 60 -----QSPSQELLFSS-----VPSRSHAHQSELDLSPLTGTSSRVEG 101
 Qy 104 A--DPCGTVTGSWPLGACAVGSCMTAMRRTRSPALPASAGSAGTDEGEDEQMIESIE 161
 Db 102 TPRSGIRGTARQRPDLGSA-----RKVKQV-----DLHSDQPAABELV 140
 Qy 162 NLEDLKGHSVRE---VSMAGPRLEIHRFNFLRTHVDSHGHN-----VPRKER 207

Db 141 TSEQSLGQKLVWGTVDNVNVA-----TCKEKFORVQVOTDPSAKEDNNVGLDNLNBIYMQR 196
 Qy 208 ISDMCKENRESLVVNYEDLAAREHVLAYFLPEAPAEALQIIDEAALEAVLAMYPKYDRIT 267
 Db 197 LEEINVGDPPNLNDCDHLRNFDDQLYRQLVCYQEVIPTEDMAANEIFFERYP--DSIL 254
 Qy 268 NH-IHVRISHLPLVEELRSRLQLHLNQLIRTSQVVTCTGLVLPOLSMVKYKNCNKNFVYG 326
 Db 255 EHQIQVRPYNALKTRNMRSLNPEDIDQLITISGVIQTSQIIPEMQEAFFCQCAFTTR 314
 Qy 327 PFCOSQNOEVKPGSCPCQAGPEVNMETIYQYQIRIQESPGKVAARRLPRSKDAI 386
 Db 315 VEI-DRGRIAPSVCKHCNTHSMALIHNRSMFSDKQMIKQLQESPEDPAQGTHTTILY 373
 Qy 387 LLADLVDSNAGDELTGTIYH-----NNYDGSNTANGFPVFATVILANHVAKONKV 440
 Db 374 GHNDLVKQVQGRVNTGIYRAVIRVNP-----RVNRKSVYKTHDVIHYRKTDSKR 428
 Qy 441 AVG-----ELTDEDVKMITSLSKQIQGIEKIFASIAPIYGHEDIKRGPALALFGE 492
 Db 429 LHGIDEDTEQKLFTEERVAMKLEAKPDIYERLAAALAPSIYEHEDIKKGILLQLFGGT 488
 Qy 493 PK--NPGGKHKVRGDIINVLLCGDPTAKSQFLKIEKVSSRAIFTTGGQASAVATYVQ 550
 Db 489 RKDFSHTGRGKFAEINILLCGDPGTSKSQLQVYVNLVPRGQVTSKGSSAVGLTAYVM 548
 Qy 551 RHPVSRWTLGAGALVLDRCVGLIDFEDKNDODRTSIHEAMBOQOSISIKAGIVTSIQ 610
 Db 549 KDPETRLQVLOTGALVLDSDNGICCIDFEDKNESTRSVLHEVMEQQVLSIAKAGIICQLN 608
 Qy 611 ARCTVIAAANPIGGYRPSLTFSENVDLTPETIISRFIDILCVVRDTPVQDMLARFVVG 670
 Db 609 ARTSVLAAANPIESQWNPCKTTIENIQLPHTLLSRFDLFLMDPQDEAYDRRLAHLVA 668
 Qy 671 SHVRHSHNSKEEGLANGSAAPAMPNTYGVPELPQEVLLKYYIYAKERVHPKLNQMDQD 730
 Db 669 LYYQSEEQAE-----LDMAVLKDYIAAHSTIMPRLSEASQ 708
 Qy 731 KVAKWYSLDRKESMATGIPITVRHIESMHSRGGPRAHPSAGLDCRRRRQHGHRP----- 785
 Db 709 ALIEAYVDMRKIGSRGMVSAYPROLES-----RRAEHAHVRSNKV 753
 Qy 786 -----DAGELHRTTEVQRHSMRK-----TFARYLSFRDNNEALLLILKOLVAEQ 831
 Db 754 ETIDVEEAKRLHREALKQSDPTGTIVDISILITGMSATARKKEELAQ-VLAKLIQSK 812
 Qy 832 -----VTYQRNRFCAQQDTIEVPEKDLVDKARQINHLNLS 866
 Db 813 GKTPALYQQLFEDLRGSDAAITKDMFDEA-----LHALA 848

RESULT 12
 T41059
 N:Alternation: licensing factor MCM3 - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
 C:Accession: T41059; T43795
 R:Murphy, L.; Harris, D.; Lyne, M.; Rajandream, M.A.; Bartell, B.G.
 submitted to the EMBL Data Library, September 1998
 A:Reference number: 221968
 A:Accession: T41059
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-879 <MUR>
 A:Cross-references: EMBL:AL031525; PIDN:CAA20668.1; GSPDB:CN00068; SPDB:SPQC1682.02c
 R:Sherman, D.A.; Forsburg, S.L.
 Nucleic Acids Res. 26, 3955-3961, 1998
 A:Title: Schizosaccharomyces pombe Mcm3p, an essential nuclear protein, associates t
 A:Reference number: 222684; MUID:98371232; PMID:9705504
 A:Accession: T43795
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-879 <SHE>

A:Cross-references: EMBL:AF063864; NID:g3139136; PIDN:AAC32263.1; PID:g3139137

C:Keywords: cell cycle control; DNA replication initiation; nucleus; phosphoprotein

C:Experimental source: strain Sp.011

C:Genetics:

A:Gene: mcm3; SPCC1682.02c

A:Map position: 3

C:Superfamily: human replication licensing factor MCM3; MCM homology

C:Keywords: cell cycle control; DNA replication initiation; nucleus; phosphoprotein

F:117-665/Domain: MCM homology <MCM>

Query Match 18.98; Score 873; DB 2; Length 879;

Best Local Similarity 34.18; Pred. No. 2.7e-47;

Matches 219; Conservative 108; Mismatches 225; Indels 90; Gaps 13;

Qy 188 FKNFLRTHVDSHGNVFKRISDWCKENRESLVVYEDLAAREHVLAYFLPEAPAELOI 247

Db 17 FOEYLEHDTDDANTLYQEAIRLMLNMQRLVINIDELRDYNRELADGVLKPLEVEP 76

Qy 248 FDEAALEVLAMYPKYDRITNHIHVRISHLPLVEEL-----RSLROLHL 291

Db 77 FDEALRWVSTLIDP-----VVHKDLKFLFYVGRSGFDHVNPRTLRAMHL 125

Qy 292 NQLIRTSVVTCTGVLPLQSLSMVKYCNKCNFVLGPPCQSONOEKPGSC-----PEC 344

Db 126 NKMISLEGIVTRCSFVRPKVIKSVHYCEATK--RHFFKQYADATMNGGLSFQSTVYPTQD 183

Qy 345 QSAGPPEVNMETIYQYQIRIQESPGKVAARLRPSKDAILLADLVDSNAGDETEL 404

Db 184 ENGNPLSIEGFTFRDQHSIQEMPERAPGOLPNSIDILLDDLDVIVKPGDRNVIV 243

Qy 405 GIYHNNYDGLSNTANGPPVFATVILANHVAKDNKVAVG-----ELTDEDVKMITSLS 457

Db 244 GQVRSM--GSKTSGNTSATERTVLANNVLLGNKPGLVGGALDITDADIRNKLA 301

Qy 458 KDOQIGEKIPASTAPSIYGHEDIKRGPALALFGEKPN--PGKHKVRGDNVLLCGDPT 516

Db 302 RKNVFEKLLSLAPSIGYEYKQAILLLLAGTEKNTLNGTH--IRGDNILMVGDPST 360

Qy 517 AKSOFKYIKBSRSRAITFGOGASAVAVTAYVQRHVPVSREWLEAGALVADRGLCID 576

Db 361 AKSOLLRFVLTAPLAIATGTRSSGVGLTAANTDKETGERLEAGAMVADRGVVCI 420

Qy 577 EFDKMDQDRTSIHEAMEQOISISKAGIVTSQARCTVIAAANPIGGRYDPSLTFSENV 636

Db 421 EFDKMSDIDRAIVEVMEQOVTIAKAGIHTSLNARCQVIAAANPIYQYDIRKDPQNI 480

Qy 637 DLTEPIISRFDILCVVDTVPQDEMLARFVVGSHVRHHPNSKEE-----EGLANG 688

Db 481 ALPDSMLSRFDLLFIVTDDIDDKRALSEHVLRMH--RYLPPGVPEGTPVRDSLSNVLN 539

Qy 689 SAAEPAMPNTYGYVEPLPOEY-----LKKYIYAKERV 720

Db 540 GATNAAGVSTENVE--QEVETPVWETTFSSLLHANARTKKELLINVFVKYIQYAKSRI 596

Qy 721 HPKLNQMDQKVKAMYSDLRKESMATG---SIPITVRHIESM 759

Db 597 HPILNQATAEYITNIYGLRNDLQGNQRRTSPLTARTLETL 638

RESULT 13

C84540

probable CDC21 protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: C84540

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.;

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: C84540

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-720 <STO>

A:Cross-references: GB:AE002093; NID:g4544386; PIDN:AAD22296.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g16440

A:Map position: 2

Query Match 18.5%; Score 858; DB 2; Length 720;

Best Local Similarity 29.28; Pred. No. 1.8e-46;

Matches 238; Conservative 132; Mismatches 264; Indels 182; Gaps 22;

Qy 133 TRSALPASASAGAGTDEGEED-----BQMIESIENLE 164

Db 3 TDEPLPSS-----DDGEEDGGDDTPTTFVMTGNISVQDVKSAIEMFVKHFRARENSD 55

Qy 165 DLKSHSVREWVWSMAGPR--LEIHHRPKNLFRLTHVDSHGHNVFKERISDMCKENRESLVN 222

Db 56 DL-----FREGKYMVSIRKVIIEGEW-----IDVDAFDV-----D 87

Qy 223 YE-DLAAREHVLAVFLPEAPAEALLOIFDEAALEVLAMYPKYDRITNHIHVRISHLPLVE 281

Db 88 YDPDLNK-----WVRYPLEVLAIEDIVLMDIVSTINRLFEE---HVQVRIENLTST 137

Qy 282 ELKSLRQLHLNQLIRTSVGVVTSCTGVLPLQSLSMVKYCNKCNFVLGPPCQSONOEKPGSC 341

Db 138 SMRNLPNDIEKMISLKMIIIRSSIIPEIREAVFRCLVCGYFSDPIIVDRGKISEPPTC 197

Qy 342 --PECQASGPEVNMETIYQYQIRIQESPGKVAARLRPSKDAILLADLVDSNAGD 399

Db 198 LKQECMTKNSMTLVHNRRCFADKQIVRLQETPDEIPEGGTPHTVLSLLHKLVDNGRPGD 257

Qy 400 EIELTGIYHNNYDGLSNTANGFP-----VEATVILANHVAKDNKVAVG----- 443

Db 258 RIEVTGIYR-----AMTVRVGPAHRTVKSVFKTYIDCLHI--KRAKLSRAEDPMDVDN 310

Qy 444 --ELTDEDVKMITSLSKDOQIGEKIFASIPSYGHEDIKRGPALALFGEKPNPGGKHK 501

Db 311 SLRRVDEDLRKFOELSKOPDIYERLSRAPNIWELDDVKKLLCQLFGGNALNASCAN 370

Qy 502 VRGDIINVLLCGDPTGAKSOFKYIKBSRSRAITFGOGASAVAVTAYVQRHVPVSREWLE 561

Db 371 FRGDIINVLLCGDPTGAKSOFKYIKBSRSRAITFGOGASAVAVTAYVQRHVPVSREWLE 430

Qy 562 AGALVLADRGVCLIDFEDKMDQDRTSIHEAMEQOISISKAGIVTSQARCTVIAAANP 621

Db 431 SGALVLSDRIGICIDEFDKMSDSARMLHEVMEQOVTIAKAGIHTSLNARTSVLACANP 490

Qy 622 IGGRYDPSLTFSENVDLTEPIISRFDILCVVDTVPQDEMLARFVVGSHVRHHPNSKE 681

Db 491 SGSRYNPRLSVNIHLPTLLSRFDLIYLIDLPDQETDRRLAKHIVALHFEAESAQ 550

Qy 682 BEGLANGSAAEPAMPNTYGYVEPLPOEVLKYYIYAKERVHPKLNQMDQKVAKMYSLRK 741

Db 551 -----EADITTLTYTVSYARKNIHPKLSDEAAEELTFGYELRK 590

Qy 742 ESMATGS---IPITVRHIESMSHGGPRAHPAGLCRRRRRHHCHPDRACELHR----- 792

Db 591 AGKFAGSKKVITATPQIESL-----IRLSALARMFSEVWEKHVDVDEAFRLRVA 643

Qy 793 -----HTEVQRHRSRMK--TFA---RYLSFRDRNNELLFLILKQLVAE 830

Db 644 MQQSATDHATGTDMDLINTGVASERMRDTPASSIRDIATALEKITMEMLL----- 695

Qy 831 QVYQYQRNFGAQDQTIIEVPEKDLVDKARQNIHNL 866

Db 696 QLEELKXKGGNINT-----EIHLDHVS 718

RESULT 14

S56766

replication licensing factor MCM4 - mouse

N:Alternate names: cell division cycle control protein CDC21/CDC54

C:Species: Mus musculus (house mouse)
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 16-Jun-2000
C:Accession: S56766
R:Kimura, H.; Takizawa, N.; Nozaki, N.; Sugimoto, K.
Nucleic Acids Res. 23, 2097-2104, 1995
A:Title: Molecular cloning of cDNA encoding mouse Cdc21 and CDC46 homologs and characterized the function of Cdc21.
A:Reference number: S56766; MUID:95334361; PMID:7610039
A:Accession: S56766
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-862 <KIM>
A:Cross-references: EMBL:D26089; NID:g940405; PIDN:BAA05082.1; PID:g940406
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, December 1993
C:Comment: The complex of six MCM proteins is one of several proteins that must be bound phosphorylated and dissociate from the chromatin.
C:Complex: The predominant form is a heterohexamer of MCM2, MCM3, MCM4, MCM5, MCM6, and C:Function:
A:Description: part of the replication licensing system that permits DNA replication to C:Superfamily: replication licensing factor MCM4; MCM homology
C:Keywords: cell cycle control; DNA replication initiation; heterohexamer; nucleus; phospho F:270-763/Domain: MCM homology <MCM>

Query Match 18.5%; Score 856; DB 1; Length 862;
Best Local Similarity 29.0%; Pred. No. 3.2e-46;
Matches 275; Conservative 140; Mismatches 331; Indels 202; Gaps 33;
2 ASSPAQR--RRG-----NDPLTSSPGRSSRRDALT-----SSPGRDL--PPFE 41
5 ASTPSRRSRRGRVPTQSLRSESSRSPNRRRGGDSSTGELLPMPTSPGADLQSPAQ 64
42 DESEGLLTGEPLEEEDEGEELIGDMERDVRAPELDAYRAEGLALDDEDEVEELT---- 97
65 N-----ALFSSPPQMHSIAIP-----LDFVSSPLTYGTP 94
98 ASRREAA--DGPCTGVTGSPGLGACAVGSCWMTAMRRTRSRALPASAASGAGTEGDEEDQ 155
95 SRVEGTSPRGVGTVPQRPDLS-----ARKGLQVDLS-----DGAAD 137
136 MISTENLEDLKGHSVREW---VSMAGPRLEIHHFRKFLRTHVD---SHGHNV----- 203
138 IVPSEQL---GOKLWIGTDVNV---JCKENFRQFLOCTDPLAKEENVGIDITQ 189
204 --FKRISDMCKENRESLVNVEDLAAREHVLAYFLPEAPAEILQIFDEALEVVLAMYP 261
190 PLYMQQLGEIITGEPFLNVAECHEIKSFKNLYROLISYPOEVIPTFDMVANEIPIFDYR 249
262 KYDRITNH--IIVRISHLPLVEELSLROLHNLNLTSGVTSCTGVLPLSMVKYCNK 320
250 --DSLEHQIQRFPFNALKTKSMRNLPEDIDOLITSGMVIRTSQILPEMQEAFQCQV 307
321 CNFVLGPPCQSONQEV-----KPGSCPCQSQAGPFVNMETIYQYQRIQRIQSPGV 374
308 C-----AHTTRVEIDRGRAEPCSCVHCHTTHSMALHNRSFSDKQMIKLOESPEDM 360
375 AARLPRSKDALLA--DLVDSNAGDEIELTGIYH-----NNDGSLNTANGPPVPAT 426
361 PAGQTPHT--IVLFAHNDLVKQPGDRVNTGTIYRAVIRNP-----RVSNVKSVYKT 413
427 VILANHVAKDKNVAVGELTDED-----VKMITSLSKDQOIGEKIFASTAPSIYG 476
414 HIDVTHYKTKDAKRLUG--LDEABQKLFSEKRVKLLKELSKKDPDIYERLASALSIYE 471
477 HEDIKRGPALALFGGEPK--NPGGKHVKGDIINVLGDPGTAKSQFLKYEKVSRAIF 534
472 HEDIKGLILLFGTGRKDESHTRGKFAEINILLGDPGTAKSQFLKYEKVSRAIF 531
535 TTGQASAVATYAYVQRPVPSREWTLEAGALVLDAGVCLIDDEFKMDNDQDRTSI 594
532 TSGKSSAVGLTAYYMKDPETRLQVLQTCALVLSNIGCCIDEDEFKMDNESTRSLVHE 591
595 QOSTISKAGIVTSLOARCTVIAAANPIGGRYDPSLTFSENVDLTFPIISRFDLVVRD 654
592 QOTLSIARAGIICQLNARTSVLAANPIESQWNPKNKTTIENIQLPHTLLSRFDLFLMLD 651

655 TVDPVQDEMLARFVYVGVSHVRHHPSNKEEGLANGSAEPAEMPTNYGVEPLPOEVLKYYII 714
652 PQDEAYDRRLAHLHLVLYVQSEEQVEEF-----LDMVVLKDYTA 691
715 YAKERVHKLQMDQDKVAKMYSDURKESMATGSPITVTHIESHSGGGPRAHPSAGLC 774
692 YAHSTIMPLRSEASQALIEAYVNNRKIGSSRGVMYSAYPRQLESLL----- 737
775 DRRRRQHGHR-----DAGELHRRHTEVORHRSMRK-----TFARYLSFRD 815
738 --RLAEAHAKVPSNKEVAIDVEEAERKRLHREALKQSATDPTGIVDISILITLTMGSATSRK 796
816 NNEILLFLIKQLVABQVNT---YQR--NRFGAQDQTIETVEPEKDLVDKA 857
797 KEELAEALRKLLSKGKTPALKYQQLFEDIRGQSDT--AITKDMFEA 842

RESULT 15

T52118

probable replication licensing factor MCM3 [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 07-Mar-2003
R:Dodeman, V.L.; Phan, T.; Sabelli, P.; Bergounioux, C.
submitted to the EMBL Data Library, September 1997
A:Description: Expression analysis of Arabidopsis thaliana MCM3 homolog during the cell cycle.
A:Reference number: Z25964
A:Accession: T52118
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-776 <DOD>
A:Cross-references: EMBL:AJ000058; PIDN:CAA03887.1
C:Superfamily: human replication licensing factor MCM3; MCM homology

Query Match 18.5%; Score 855; DB 2; Length 776;
Best Local Similarity 32.6%; Pred. No. 3.2e-46;
Matches 232; Conservative 117; Mismatches 283; Indels 80; Gaps 15;

181 RLEIHHRRKFNELTHVDSHGHNVEKERSDMCKENRESLVNVEDLAAREHVLAYFLPEA 240
8 RLHRKDRFIQFL-----DSMYEELKALVHQRHRLIINISDIIHHFREVASRLKN 59
241 PAELQIFDEALEVVLAMYPKYDRITNHHVRIIS--HLPLVEELSLRQLHNLIRTS 298
60 PNEYMQSCDAATEATRAIDPKYKLEGEVLVGFEGYFVSVVTPRELLSDFIGSMVCVE 119
239 GWTTSCTGVLQSLSMVKYCNKCNFVLGPFQSQONQEVKPGSCPCQSQAGPFVNMETI 358
120 GIVTKCSLVRPKVKSVMHFCS-----TGFTNRDRIITSHAGLPTGVSYYTRDDKGNLL 175
359 -----YQYQRIQIESPGKVAARLPRSKDAILLADLVDSNAGDEIELTGIYHNN 410
176 VTEYGLCKYKHQTLISQIEVPEANAAPGQLPRSDVIAEDDLVDSCNPGDRVSVFGIYKAL 235
411 YDGSUNTAGPFVATVILANHVAKDKNVAVGELTDEDVKMITSLSKDQOIGEKIFASI 470
236 PGKSGSVNG--VFRTILANNIALLNKEANPIVTKQDLNKNKIARRDDAFOLLARSL 293
471 APSIYGHEDIKRGPALALFGGEPKN--PGGKHVKGDIINVLGDPGTAKSQFLKYEKVS 529
294 APSIYGHAWIKAVVLLMLGGVEKRLKNGTH--LRGIDNMMVMGDPDSVAKSOLLRAIMNIA 352
530 SRAIFTTGGQASAVATYAYVQRPVPSREWTLEAGALVLDAGVCLIDDEFKMDNDQDRTSI 589
353 PLAISTTGGSGVGLTAATVSDQETGERLERLEAGAMVLADKIVCIDDEFKMDNDQDRAI 412
590 HEAMDDQSTISKAGIVTSLOARCTVIAAANPIGGRYDPSLTFSENVDLTFPIISRFDL 649
413 HEVMEQQTVTAKAGIHASLNARCSVAAANPIYDRLSLPTKKNIGLPSLLSRFDLL 472
650 CVVRTVDPVQDEMLARFVYVGVSHVRHHPSNKEEG-----LANGSAEPAEMPTNYGV-- 701

Db 473 FIVLDQMDAGIDSMISEHVLRMH--RYKNDRSEAGPDGSLPYAREDNAESEMFKYNQTL 530
QY 702 -----EPLPOEVLKYYIYAKERVHPKLNQMDQDKVAKMYSDLRKESMAT---G 747
Db 531 HGKKKRGQTHDKTLTIKFLKYYIYAKHRITPKLTDEASERIAEAYADLERNAGSDTKTGG 590
QY 748 SIPITVVRHIESMSHGGPRAHPSAGLCDRRRROHG-----HPRDAGELHRHTEV 796
Db 591 TLPITARTLETIIRLA--TAHAKMKLSSEVTKADAEAAALKMLMNFATYHQELTEMDREQE 648
QY 797 QRHRSRMRKTFARYLSFRONNELLFILKQLVAEQVYQNRNRFCAOODTIEV 848
Db 649 EROREQAEQ-ERTPSGRGNQR-----RNNEDGAENDTANV 683

Search completed: September 12, 2003, 13:14:22
Job time : 31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2003, 13:09:25 ; Search time 16 Seconds
(without alignments)
2621.739 Million cell updates/sec

Title: CAA47749

Perfect score: 4630

Sequence: 1 MASSPAQRRRGNDPLTSSPG.....LFRNKFSDLRKRMILQOF 892

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	DB ID	Description
1	4078.5	88.1	895	1	MCM2_HUMAN
2	3941.1	85.1	904	1	MCM2_MOUSE
3	3539.5	76.4	886	1	MCM2_XENLA
4	2649.7	57.2	887	1	MCM2_DROME
5	1951.5	42.1	830	1	CC19_SCHPO
6	1812.6	39.1	868	1	MCM2_YEAST
7	877.1	18.9	863	1	MCM4_HUMAN
8	874.5	18.9	863	1	MCM4_XENLA
9	873.1	18.9	879	1	MCM3_SCHPO
10	856.1	18.5	862	1	MCM4_MOUSE
11	848.5	18.3	933	1	CC54_YEAST
12	843.1	18.2	931	1	CC21_SCHPO
13	840.1	18.1	821	1	MCM6_MOUSE
14	835.1	18.0	821	1	MCM6_HUMAN
15	825.5	17.6	866	1	MCM5_XENLA
16	816.5	17.8	735	1	MCM4_DROME
17	802.1	17.3	971	1	MCM3_YEAST
18	793.5	17.1	1017	1	MCM6_YEAST
19	789.5	17.1	734	1	MCM5_HUMAN
20	789.1	17.0	775	1	MCM5_YEAST
21	786.5	17.0	892	1	MIS5_SCHPO
22	783.1	16.9	720	1	MCM7_XENLA
23	781.1	16.9	719	1	MCM7_HUMAN
24	778.5	16.8	808	1	MCM3_HUMAN
25	777.5	16.8	716	1	PROL_ARATH
26	776.5	16.8	719	1	MCM7_MOUSE
27	772.5	16.7	720	1	NDA4_SCHPO
28	770.1	16.6	760	1	MCM7_SCHPO
29	769.1	16.6	807	1	MCM3_XENLA
30	768.1	16.6	733	1	MCM5_MOUSE
31	764.5	16.5	812	1	MCM3_MOUSE
32	761.1	16.4	759	1	MCM5_CAEEL
33	758.5	16.4	840	1	MCM8_HUMAN

RESULT 1

ID	MCM2_HUMAN	STANDARD	PRT	895 AA
AC	P49736	Q14577; Q15023; Q969W7; Q96AE1; Q9BRM7;		
DT	01-OCT-1996	(Rel. 34, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	15-SEP-2003	(Rel. 42, Last annotation update)		
DE	DNA replication licensing factor MCM2 (Nuclear protein BM28).			
GN	MCM2 OR CDC11 OR BM28 OR KIAA0030.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94230605; PubMed=8175912;			
RA	Todorov I.T., Pepparkok R., Philipova R.N., Kearsey S.E., Ansorge W., Werner D.;			
RT	"A human nuclear protein with sequence homology to a family of early S phase proteins is required for entry into S phase and for cell division."			
RL	J. Cell Sci. 107:253-265(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96051387; PubMed=7584026;			
RA	Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y., Sato S., Nagase T., Seki N., Ishikawa K.-I., Tabata S.;			
RT	"Prediction of the coding sequences of unidentified human genes. I. The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by analysis of randomly sampled cDNA clones from human immature myeloid cell line KG-1."			
RL	DNA Res. 1:27-35(1994).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Mimura S., Nishimoto S., Kubota Y., Takisawa H., Nojima H.;			
RL	Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUP=Brain, Lung, Lymph, Muscle, Placenta, and Uterus;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,			

ALIGNMENTS

34	751	16.2	810	1	MCM6_CAEEL
35	748.5	16.2	600	1	MCM3_MAIZE
36	748	16.2	845	1	CC47_YEAST
37	650	14.0	507	1	MCM6_RAT
38	644	13.9	597	1	MCM3_ENTHI
39	591	12.8	495	1	MCM8_MOUSE
40	466.5	10.1	682	1	YE89_METJA
41	442	9.5	762	1	Y961_METJA
42	430	9.3	759	1	Y363_METJA
43	271	5.9	602	1	YZ13_METJA
44	154.5	3.3	516	1	Y1FB_ECOLI
45	137.5	3.0	506	1	Y1FB_SALTY

P34647	caenorhabdi
Q43704	zea mays (m
P38132	saccharomyc
Q22724	rattus norv
Q24849	entamoeba h
Q9cww1	mus musculu
Q58884	methanococc
Q58371	methanococc
Q57809	methanococc
Q50275	methanococc
P22787	escherichia
P37015	salmonella

SEQUENCE FROM N.A.
RX MEDLINE=94129084; PubMed=8298187;
RA Miyake S., Okishio N., Samejima I., Hiraoka Y., Toda T.,
RA Satoh I., Yanagida M.,
RA "Fission yeast genes nda1+ and nda4+, mutations of which lead to
RT S-phase block, chromatin alteration and Cdz+ suppression, are members
RT of the CDC46/MCM2 family.";
RL Mol. Biol. Cell 4:1003-1015(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman C.M.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Coffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT The genome sequence of *Schizosaccharomyces pombe*.;
RL Nature 415:871-880(2002).
CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN DNA REPLICATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE MCM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U08048; AAC48930.1; -;
DR EMBL; S68468; AAC60569.1; -;
DR EMBL; AL121863; CAB58403.1; -;
DR PIR; B48723; B48723.
DR GeneDB_Spomb; SPBC4.04c; -;
DR InterPro; IPR001208; MCM.
DR Pfam; PF00493; MCM; 1.
DR ProDom; PD001041; MCM; 1.
DR SMART; SM00350; MCM; 1.
DR PROSITE; PS00847; MCM_1; 1.
DR PROSITE; PS50051; MCM_2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW DNA replication; Zinc-finger; ATP-binding; Cell cycle.
FT ZN_FING 334 360
FT DOMAIN 484 691
FT C4-TYPE (POTENTIAL).
FT NP_BIND 534 541
FT ATP (POTENTIAL).
SQ SEQUENCE 830 AA; 92831 MW; A5BF0AA29661BD66 CRC64;
Query Match
Best Local Similarity 42.1%; Score 1951.5; DB 1; Length 830;
Matches 418; Conservative 116; Mismatches 238; Indels 95; Gaps 12;
QY 1 MASSPAQRRRNDPLTSSPGRRRTDALTSPPGRDLPPFDESEGLLTGEGPLE----- 55

Db 1 MDSFRKGRGRDSESLPFSESENLSLGATPLSPSPSPPEFSDAEAL-VSEEDIEDLGE 59
QY 56 ---EEDGEELIGDMERDYRAIPELDAYAEGLADLDEVEELTASRRRAADGPCGT 111
Db 60 ALDVEDEGEELFEGMERDYOQNELDYDIELD-DNDLEELDIGARRAVD----- 112
QY 112 TGSWPGLCACAVGSCMTAMRRTRSALPASAA-----SGAGT----- 147
Db 113 -----AFLRRDIELDAAAGRTKPAFLQDEDDDDLSNLGTGTRHRH 156
QY 148 -----EDGEEDQMIIESINLEDLKCHSVREWSMAGPRLEIHRHFKNELRTHVDSHG 200
Db 157 IYDEYSPNVALDESSELPLESIADVKADSAEWVTLDPVVRTIAREFKNELLEYDENG 216
QY 201 HNVKERISDMCKENRESLVNVEDLAAREHVLAYELPEAPAEALLOIIDEAALVVLAMY 260
Db 217 TSVYGNRIITLGEVNAESLMVNYAHLGSKPILAFLANAPAFIRIDRVALNEATLUHY 276
QY 261 KYDRITNHIHVRISHLPVLEELRSLRQLHNLQIRTSVYVTSCTGVLPLQSLSMVKYCNK 320
Db 277 PDYERIHSDIHVRITNLTCTFLRDLRQSHLNLVRSVGVVTRTGLFPQLKYIRFTCTK 336
QY 321 CNFVLGPPCQSQNOEVKPGSCPECOSAGPVEVNMETIYQNYQRIQESPGKVAARLPL 380
Db 337 CGATLGPFPFODSSVEVKISFCHNGSSRGPPFVINSERTVYNNYQRIITLQESPTVPSGRLP 396
QY 381 RSKDAILLADLVDSNAGDEITELGIYHNNDGSLTANGPVPVATVILANHVAK----- 435
Db 397 RHREVVILLADLVAKPGEEIDVTGIYRNNDASLNTKNGPVPVATVILANHVAK----- 456
QY 436 -KDNKAVAGELTDEDVKMITSLSKDDQIGERIFASIAPISYGHEDIKRGPALALFGGEPK 494
Db 457 NTDDDFSLSLTDDDEERIRALAKSPDINHRIASMAPSIYGHRSIKTAIAAALFGGVPK 516
QY 495 NPGKHVKRGDINVLLCGDPTAKSQFLKYIEKVSRAIFTTGGQASAVATAYQRRPVP 554
Db 517 NINGKHKIRGDNVLLGDPCTAKSQFLKYIEKTAHRAVAVFATGGQASAVGLTASVRKDKPI 576
QY 555 SREWTLEAGALVLADRGVCLIDFDMKDQDRTSHEAMEQOSISISKAGIVTSLQARCT 614
Db 577 TNEWTLEAGALVLADRGVCLIDFDMKDQDRTSHEAMEQOSISISKAGIVTSLQARCT 636
QY 615 VIAANPIGGYDPSLTFSENVDLTEPIISRFIDLCVVVRDTPVQVDEMLARFVVGSHVR 674
Db 637 IIAANPIGGYNTTIPFNQNVLTETILSRFDILQVVKDTPVNPEDQLANFVVSISHIR 696
QY 675 HPSNKEEGLANGSAAPAM-----PNTYGV--PLPOEVLLKYYIYAKERVHPKLNQ 726
Db 697 SHP-----AFDPNMDVLKVPETGTGIDAKPIPDLLLRKYTHFAREKVPRLQQ 744
QY 727 MDQKVAKWYSDLRKESMATGIPITVRHIESMHHGGPRAHPSAGLCDRRRRQHGHPRD 786
Db 745 MDEEKISRLYSDMRRESLATGSPYTVRHLESATRLS--EAFAKMQLSEFVRPSHIDKAI 802
QY 787 AGEIHRHTEVQR---HRSMRKTFARYL 810
Db 803 QVIIDSFVNAQKMSVKRSLSRTFAKYL 829

RESULT 6

MCM2_YEAST
ID MCM2_YEAST PRT: 868 AA.
AC P29469;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA replication licensing factor MCM2 (Minichromosome maintenance
DE protein 2).
GN MCM2 OR YBL023C OR YBL0438.
OS *Saccharomyces cerevisiae* (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; *Saccharomycotina*; *Saccharomycetes*;
OC *Saccharomycetales*; *Saccharomycetaceae*; *Saccharomycetes*.

DE DNA replication licensing factor MCM4 (CDC21 homolog) (P1-CDC21).

GN MCM4 OR CDC21.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RP [1]

RN SEQUENCE FROM N.A.

RA Hu B.;

RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 1-712 FROM N.A.

RX MEDLINE=97430835; PubMed=9465298;

RA Connelly M.A., Zhang H., Kieleczawa J., Anderson C.W.;

RT "The promoters for human DNA-PKcs (PRKDC) and MCM4: divergently

RT transcribed genes located at chromosome 8 band q11.3";

RL Genomics 47:71-83(1998).

RN [3]

RP SEQUENCE OF 1-23 FROM N.A.

RX MEDLINE=97430835; PubMed=9284934;

RA Ladenburger E.M., Fackelmayr F.O., Hameister H., Knippers R.;

RT "MCM4 and PRKDC, human genes encoding proteins MCM4 and DNA-PKcs, are

RT close neighbours located on chromosome 8q12--q13.3";

RL Cytogenet. Cell Genet. 77:268-270(1997).

RN [4]

RP SEQUENCE OF 440-863 FROM N.A.

RC TISSUE=Cervix;

RX MEDLINE=94089373; PubMed=8265339;

RA Hu B., Burkhardt R., Schulte D., Musahl C., Knippers R.;

RT "The P1 family: a new class of nuclear mammalian proteins related to

RT the yeast Mcm replication proteins";

RL Nucleic Acids Res. 21:5289-5293(1993).

CC -!- FUNCTION: INVOLVED IN THE CONTROL OF DNA REPLICATION.

CC -!- SUBCELLULAR LOCATION: Nuclear (BY similarity).

CC -!- SIMILARITY: BELONGS TO THE MCM FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

DR EMBL; X74794; CAA52801.1; -;

DR EMBL; U63630; AAC52018.1; -;

DR EMBL; U90415; AAB51723.3; -;

DR PIR; S65954; S65954.

DR Genbank; HGNC:6947; MCM4.

DR GK; P33991; -;

DR MIM; 602638; -;

DR GO; GO:0005634; C:nucleus; NAS.

DR GO; GO:0005524; F:ATP binding activity; NAS.

DR GO; GO:0006260; P:DNA replication; NAS.

DR InterPro; IPR003593; AAA_Attrase.

DR InterPro; IPR001208; MCM.

DR Pfam; PF00493; MCM; 1.

DR ProDom; PD01041; MCM; 1.

DR SMART; SM00382; AAA; 1.

DR SMART; SM00350; MCM; 1.

DR PROSITE; PS00847; MCM_1; 1.

DR PROSITE; PS00501; MCM_2; 1.

KW Transcription regulation; DNA-binding; Nuclear protein;

KW DNA replication; ATP-binding.

FT DOMAIN 458 667 MCM.

FT NP_BIND 510 517 ATP (POTENTIAL).

FT CONFLICT 62 206 T -> P (IN REF. 2).

FT CONFLICT 206 206 Q -> P (IN REF. 2).

FT CONFLICT 650 650 M -> L (IN REF. 2).

SQ SEQUENCE 863 AA; 96610 MW; 50010575C41BA90C CRC64;

Query Match 18.9%; Score 877; DB 1; Length 863;

Best Local Similarity 30.4%; Pred. No. 1.3e-47;

Matches 246; Conservative 129; Mismatches 280; Indels 154; Gaps 23;

Qy 3 SSPAORRRGND-----PLTSSPG---RSSRRFDALTSSPGRDLPPFESEGLLGT 51

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

31 SSPSQRREGDSTSTGELQPMPTSPGVDLQSTAQDVLFFSSP----- 72

Qy 52 GPLEEEDGEGELIGDMERYRAIPELDAYEAGLALDDEVEELT-----ASRREA--D 105

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

73 -----PQM---HSSAIPDLDFVSSPLTYGTPTPSRVEGTPRS 105

Qy 106 GPCGTVTGSPGLGACAVGSCMTAMRTRSPALPASAASGACTGDEDEQMIENLEND 165

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

106 GVRGTPVRQRPDLGLSAQGLQVDLQ-----SDGAAAEIVASEQSL----- 146

Qy 166 LKHSVREW---VSMAGPRLEIHHRFKNFLRTHVD---SHGHNV-----FKERISDM 211

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

147 --GOKLIVGTDVNVAAK-----ENQRFQRFQFDPLAKEEENGVGIDITELYQWRLEI 200

Qy 212 CKNRESLVNYEDLAAREHVLAYFLPEAPAEALQIFDEAALEAVLVWAMYKYDRITNH-I 270

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

201 NVIGEQLNVNCEHIKSFKNLYRQLISYPOEVIPTFDMAVNEIFFDRYP--DSILEHOI 258

Qy 271 HVRISHLPLVEELSLRLQLHLNOLIRTSVVVTSCTGVLQPLSMVYKNCNKFVLGPQC 330

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

259 QVRPFNAKTKNMRNLNPEIDQLITISGMVIRTSQLIPEMQEAFFQCVCAHTT-RVEM 317

Qy 331 SONQEVKPGSCPECQSAQPFVNMEEITYQYQIRIQESPGKVAARLRPSKDAILLA- 389

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

318 DRGRIAPSVCGRCHTHSMALIHNSLFSQKQIKLOESPEDMPAGOTPT--VILFAH 375

Qy 390 -DLVDSNAGDEIELTGIYH-----NNYDGLSNTANGFPVFATVILANHVAKDKNVAV 442

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

376 NDLVDKVPQGDVRNVNTGIYRAVPIRVNP-----RVSNVKSVMYKTHIDVHYRKTDAKRLH 430

Qy 443 GELTDED-----VKMITSLSKQOIGCEKIFASTAPSIYGHEDIKRGPALALFGE 492

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

431 G--LDEAEAKLFSEKRVELLKSLRPDIYERLASALAPSIYEHEDIKKGILLQFGGT 488

Qy 493 PK--NPGKHKVRGDIINVLCGDPGTAKSQFLKVIKVSRAIFTTGGASAVAVATYVQ 550

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

489 RKDPSHTGRGKFAEINILLGCDPGTSKQLQVYVNLVPRQYTSKGSSTGSSAVGLTAYVM 548

Qy 551 RHPVSREWTLBAGLVLAADRGVCLIDFEDKNDQDRTSIHEAMQOQSISIKAGIVTSIQ 610

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

549 KDPETRLQLVLOTGALVLSNDGICCTDEFKDNNESTRSVLHEVMEQOQLSIKAGITICQLN 608

Qy 611 ARCTVIAAANPIGGRYDPSLTFSENVDLTETPIISREDILCVVTRDVPQVDEMLARFVVG 670

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

609 ARTSVLAAANPIESQWNPCKTTIENIQLPHTLLSRFDLIFLMLDPQDAYDRRLAHLVA 668

Qy 671 SHVRHPSNKEEGLANGSAAEPAMPNTYGYVEPLPQEVKKYIIYAKERVHPKLNQMDQD 730

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

669 LYOSEEAQAE-----ELLDMAVLKDYIAYAHSTIMPRLSEASQ 708

Qy 731 KVAKMYSDLRKESMATGSIPTVRIHESM 759

Db |||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

709 ALIEAYVDMRKIGSSRGVMVSAYRQLES 737

RESULT 8

MCM4_XENLA STANDARD; PRT; 863 AA.

ID MCM4_XENLA

AC P30664; Q91679;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE DNA replication licensing factor MCM4 (CDC21 homolog) (P1-CDC21)

DE (X.MCM4).

GN MCM4.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;
 RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96183193; PubMed=8605878;
 RA Coue M., Kearsley S.E., Mechali M.;
 RT "Chromatin binding, nuclear localization and phosphorylation of
 RT Xenopus cdc21 are cell-cycle dependent and associated with the
 RT control of initiation of DNA replication.";
 RL EMBO J. 15:1085-1097(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=97057222; PubMed=8901561;
 RX Hendrickson M., Madine M., Dalton S., Gautier J.;
 RA "Phosphorylation of MCM4 by cdc2 protein kinase inhibits the activity
 RT of the minichromosome maintenance complex.";
 RT Proc. Natl. Acad. Sci. U.S.A. 93:12223-12228(1996).
 RN [3]
 RP SEQUENCE OF 513-588 FROM N.A.
 RX MEDLINE=93087163; PubMed=1454522;
 RA Coxon A., Maundrell K., Kearsley S.E.;
 RT "Fission yeast cdc21+ belongs to a family of proteins involved in an
 RT early step of chromosome replication.";
 RL Nucleic Acids Res. 20:5571-5577(1992).
 CC -1- FUNCTION: INVOLVED IN THE CONTROL OF DNA REPLICATION.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE MCM FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U29178; AAB01680.1; -;
 DR EMBL; U46131; AAA91232.1; -;
 DR EMBL; Z15033; CAA78751.1; -;
 DR PIR; S64720; S64720.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001208; MCM.
 DR Pfam; PF00493; MCM, 1.
 DR ProDom; PD001041; MCM; 1.
 DR SMART; SM00382; AAA; 1.
 DR SMART; SM00350; MCM; 1.
 DR PROSITE; PS00847; MCM.1; 1.
 DR PROSITE; PS50051; MCM.2; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein;
 KW DNA replication; ATP-binding.
 FT DOMAIN 458 667 MCM.
 FT NP_BIND 510 517 ATP (POTENTIAL).
 FT CONFLICT 73 73 V -> A (IN REF. 2).
 FT CONFLICT 740 741 RA -> LS (IN REF. 2).
 SQ SEQUENCE 863 AA; 97178 MW; 9A47003097EEDEZE CRC64;

 Query Match 18.9%; Score 874.5; DB 1; Length 863;
 Best Local Similarity 28.1%; Pred. No. 1.9e-47;
 Matches 264; Conservative 145; Mismatches 358; Indels 173; Gaps 25;

 QY 3 SSPAQR-----RNDPLT-----SSPGRSSRRTDALTS-----SPGRDLPFDE 43
 |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:|
 6 STPSRRNRKGRGNPPTGHGEVQSPSPRRRTDSTSGIGLLPMPPTSPGDV----- 59

 QY 44 SEGLLTGEPLLEEGDEELIGDMERYAIPELDAYEAGLADDEVEELTASREA 103
 |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:|
 60 -----SSPQGLLFSS-----VPSRHSQHSELDLSPLTYGTPSSRVGE 101

 QY 104 A--DGCPTGTGWPVGICAGVSCMTAMRRTRSRALPASASGAGTDEGEDEQMTESIE 161
 |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:|
 102 TPRSIRGTPTARQPDLSA-----RKVKQV-----DLHSDQPAEELV 140

 QY 162 NLEDLKGHSVREW---VSMAGPRLTHHRFKNFLRTHVDSHGHN-----VFKER 207
 |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:| |:|:|:|

Db 141 TSEQSLGQKLVIMGTDVNVNVA---TCKEKFRQRFVQFIDPSAKEEDNVGLDLNEPIYMR 196
 QY 208 ISDMCKENRESLVVNYVEDLAAREHVLAYELPEAPALLOIFDEAALEAVLWLPKPYDRIT 267
 Db 197 LEEINVVGDFFLNIDCDHLNRFNODLYRQLVCPQVEIPTFDMAANEIIFERIP--DSIL 254
 QY 268 NH-IHVRISHPLVEELRSLQLHLNQLIRTSQVVTSCGTGTPOLSMVKYNCKNCFVLG 326
 Db 255 EHQIQVRYPYNALKTRNMRSLNPEDIDQLITISGMVIRTSQIIPEMQEAPKQCVCAFTTR 314
 QY 327 PFCQSONQEVKPGSCPECSAGPEFVNMEETIYQNTQRIQIESPKVVAARRLPKSKDAI 386
 Db 315 VEI-DRGRTAEPVSVCKHCNTTHSMALIHNRMSFSDKQMIKLOESPEDMPAGTPHTILY 373
 QY 387 LLADLVDSNAGDEIELTGIYH-----NNDGSLNTANGFPVFATVILANHVAKDKNV 440
 Db 374 GHNDLVKVPQGDVRVNVVTGIYRAVPIRVNP-----RVNRVKSVMYKTHIDVIRKTKDSKR 428
 QY 441 AVG-----ELTDEDVKMITSLSKDDQIGEKIFASIAPIYSIYGHEDIKRGPALALFGE 492
 Db 429 LHGIDEDTEQKLFTEERVAMKLKELAAKPDIIYERLAAALAPSIYEHEDIKKGILLQFGGT 488
 QY 493 PK--NPGGKHVKRGDINVLGDPGTAQSOFKYIEKVSRAIFTTGOGASAVATYVQ 550
 Db 489 RKDFSHTGRGKFAEYVNLICGDPGTSKQLQYVFNLPVPRGQYTSKGSSAYGLTAYVM 548
 QY 551 RHPVSRWLTLEAGALVADRGVCLIDFEDKMDNDRSTSIEAMEQOQISISKAGIVTSIQ 610
 Db 549 KDPETQLVLQTLGALVLSONGICCIDFEDKMNSTRSVLHEVMEQQTLSAKAGIICQLN 608
 QY 611 ARCTVIAAANPIGRGDPDSLTFSENYDLTEPIISRDILCVVRDTPVQVDENLARFVVG 670
 Db 609 ARTSVLAAANPVESQNNPKKTTENIQLPHTLLSRDLFLMLDPQDEAYDRRLAHLVA 668
 QY 671 SHVRHPSNKEEGLANGSAAEPAMNTYGVPELVKYLKYLIIYAKERVHPKLNQMDQD 730
 Db 669 LYYQSEEQMKKEH-----LDMAVLKDYIAYARTYVYNPHLSBEASQ 708
 QY 731 KVAKMYSDLRKESMATGSIPTVVRHIESMSHGGRAPHSAGLCDDRRRRROGHPR----- 785
 Db 709 ALIEAYVSMRKIGSGRGVMSAYPROLESLL-----RRAEAHAKVRFSNKV 753
 QY 786 -----DAGELHRHTEVQRHRSRK-----TFARYLSFRDRNNELLFLILKQLVAEQ 831
 Db 754 ETIDVEEAKRLHREALKQSATDPTGTGIVDVISILTTGMSATARKKELEA-VLKLKIQSK 812
 QY 832 -----VTYQNRFGAODTIEVPEKDLVDYKARINHNLS 866
 Db 813 GKTPALKYQQLFEDLRGQSDAAITKDMFDEA----LHALA 848

 RESULT 9
 MCM3_SCHPO STANDARD; PRT; 879 AA.
 AC P30666;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA replication licensing factor mcm3 (Minichromosome maintenance
 DE protein 3).
 GN MCM3 OR SPCC1682.02C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SP011;
 RA MEDLINE=98371232; PubMed=9705504;
 RX Sherman D.A., Forsburg S.L.;
 RT "Schizosaccharomycetes pombe Mcm3p, an essential nuclear protein,
 RT associates tightly with Nda4p (Mcm5p).";

RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: REQUIRED FOR S PHASE EXECUTION.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE MCM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X58824; CAA41628.1; -;
DR EMBL; AL109957; CAB53089.1; -;
DR EMBL; AL157991; CAB76210.1; -;
DR GeneDB_SPombe: SPCC16A11.17; -;
DR InterPro: IPR001208; MCM.
DR Pfam: PF00493; MCM; 1.
DR ProDom: PD001041; MCM; 1.
DR SMART: SM00350; MCM; 1.
DR PROSITE: PS00847; MCM; 1.
DR PROSITE: PS50051; MCM; 2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW DNA replication; ATP-binding.
FT DOMAIN 493 702 MCM.
FT NP_BIND 545 552 ATP (POTENTIAL).
FT CONFLICT 872 891 MISSING (IN REF. 3).
FT SEQUENCE 931 AA; 103728 MW; BCA9B045FC62811D CRC64;
Query Match 18.2%; Score 843; DB 1; Length 931;
Best Local Similarity 28.5%; Pred. No. 2e-45;
Matches 236; Conservative 164; Mismatches 289; Indels 140; Gaps 28;
QY 2 ASSPAQRNRGNPL---TSSPGR---SSRRTDALT-----SSPGRLPPPEDES 44
DB 19 ANSSREAVDSPLFPFASPSPGSTRLTTPRTTARTPLASPLLFESSPGPNIP--QSSR 76
QY 45 EGLLGTGPLEEEDGEEELIGDGMERYDRAIPELDPAEAGLALDDDEVEELTASREAA 104
DB 77 SHLLSQNDL-----FLDSSQRTPRSTRGDHSS-----VQMSTPSRREV 119
QY 105 DPGCGVTGSGPGLGACVSGSMTAMRRTRTSALPASAASGAGTGEDDEQETESTENLE 164
DB 120 D-----PQPCGV-----TPSSLLFGSDALTFSAQPSSEVADTVRVI 159
QY 165 DLKGHSVREWS-----MAGPLRIEIHFRKFNLRTHVDSHGHNVFKEISDMCKENRESL 219
DB 160 WGTNVSIQSIASFRGLFGKKYRPEYRNEIMPPPDAB-QLVYTEALRNRMIMGLEIL 218
QY 220 VVNYEDLAAREHV-----LAYLPEAPAEALLOIFDEAALLEVVLAMT---PKYDRITNHIH 271
DB 219 NLDVQDL---KHPPTPKLYHQLYSYQELIIPINDQTIKDVMLDLTGTPPED-VLNDIE 274
QY 272 VRISHL-PL-VEELRSLRLQH---LNQLIRTSQVVTSCVTGLPQLSMVKYKNCNKFNVLG 326
DB 275 LKTIKIRPENLEKICNNRDLNPGDIDKLIISIGLVLRCTPVIPDMQAFFRCVSGC---- 330
QY 327 PFCQS-----QNEVKGCSCPE--CQSAGPFVFNMEETIYQNTQIRIQESPGKVAARRLP 380
DB 331 -HCVTVEIDRGRIAEIKPCPREVCGATNAMQLIHNRSEFADKQVLIKQETPDVVPDQGP 389

QY 381 RSKDAILLADLVDSNAGDEIELTGIYH---NNYDGSNTANGFPVPFATVILANHVAKDD 437
DB 390 HVSVLCVYDELVDLSARAGDRIVETGIFRCVPLRNPMTVKV--LFTKYVDVHVHIKQD 447
QY 438 NK-----VA-----VGLTDDVKMTSLSKDQOQGEKIFASIASIYV 476
DB 448 KRLRGDTPSTLESIDAEALQIDEVKRTISDEVEKIQVQSRDDIYDILSKSLAPSIYE 507
QY 477 HEDIKRGPALALFGEPEK--NPGGKHVGRGDNVLLCGDPGTAKSOFKAYIEKVSSRAIF 534
DB 508 MDVVRKGLLLQFLFGGTTKSFHGASPRYRGDINILMCGDPSTSKSOLLVYVHKIAPGVY 567
QY 535 TTGQGSASAVATYVQVRHPVSWREWTLEAGALVLADRGVCLIDFEDKMNDDQRTSIHEAME 594
DB 568 TSGKGSASVGLTAYITRDQDTKQLVLESGALVLDGGICCIDFEDKMSDATRSILHEVME 627
QY 595 QQSISISKAGIVTSLOARCTVTAANPIGGRVDPSTLTFSENVDLTPEIISRFDILCVVRD 654
DB 628 QQTVTYAKAGIITTLNARTSILASANPIGSKYNDLPVTKNIDLPPTLLSRFDLVYLLD 687
QY 655 TVDPVQDEMLARFVVGSHVRHHPSPNKEEGLANGSAAEPAMPNTYGVPELPQVLKKYII 714
DB 688 RVEYELDRKLANHIIVSMYMEDTPEHATD-----MEVFSVZFLSYIT 729
QY 715 YAKERVHPKLNQMDQDKVAKMYSDLRK--ESMATGS--IPITVVRHIESM 759
DB 730 YARNINNPVISEEAAKELVNAVYGMKLGEDVRASEKRITATTATTOLES 778
RESULT 13
MCM6_MOUSE
ID MCM6_MOUSE STANDARD; PRT; 821 AA.
AC P97311;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE DNA replication licensing factor MCM6 (Mis5 homolog).
GN MCM6 OR MCM6 OR MIS5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=97224213; PubMed=9077461;
RA Kimura H., Ohtomo T., Yamaguchi M., Ishii A., Sugimoto K.;
RT "Mouse MCM proteins: complex formation and transportation to the
RT nucleus.";
RL Genes Cells 1:977-993(1996).
CC -!- FUNCTION: MAY BE INVOLVED IN THE CONTROL OF A SINGLE ROUND OF DNA
CC REPLICATION DURING S PHASE. BINDS TO CHROMATIN DURING G1 AND
CC DETACH FROM IT DURING S PHASE AS IF IT LICENSES THE CHROMATIN TO
CC REPLICATE.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE MCM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D86726; BAA13159.1; -;
DR MGD; MGI:1298227; Mcm6.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro: IPR001208; MCM.
DR Pfam: PF00493; MCM; 1.
DR ProDom: PD001041; MCM; 1.
DR SMART; SM00350; MCM; 1.
DR PROSITE; PS00847; MCM; 1.

DR PROSITE; PS50051; MCM_2; 1.
KW DNA replication; ATP-binding; Nuclear protein;
FT DOMAIN 346 553 MCM.
ET NP_BIND 396 403 ATP (POTENTIAL).
SQ SEQUENCE 821 AA; 92867 MW; 79E753D2DA218CBF CRC64;

Query Match 18.1%; Score 840; DB 1; Length 821;
Best Local Similarity 27.3%; Pred. No. 2.6e-45;
Matches 234; Conservative 154; Mismatches 286; Indels 184; Gaps 21;

QY 137 LPASAAAGCTDEGEDEQMTESLENLEDLKGHSVREWVSMAGPRLEIHHRFKFLRTHV 196
DQ 137 LAAAEAGAGSQHPEVDEVAEKCKL-----FLDFLEEFQ 38
QY 197 DSHGHNFKERISDMCKENRESLVVYEDLAAREHVLAYFLPEA-----PAELLOTFD 249
DQ 197 DSHGHNFKERISDMCKENRESLVVYEDLAAREHVLAYFLPEA-----PAELLOTFD 249
QY 39 GSDSEIKYLAELIRPERNLVVSFADLFQFNQQLSTTQEEFYRYVYPLCALRALTIV 98
DQ 39 GSDSEIKYLAELIRPERNLVVSFADLFQFNQQLSTTQEEFYRYVYPLCALRALTIV 98
QY 250 EAALEVLVAMYPKYDRITNHHRISHLPLVEELSLRQLHLNQLIRTSQVVTCTGYLP 309
DQ 250 EAALEVLVAMYPKYDRITNHHRISHLPLVEELSLRQLHLNQLIRTSQVVTCTGYLP 309
QY 99 KDRKEIPPA-----KDFYVAFQDLPTRHKIRELTSSRIGLLTRISGOVVRTHPVHP 149
DQ 99 KDRKEIPPA-----KDFYVAFQDLPTRHKIRELTSSRIGLLTRISGOVVRTHPVHP 149
QY 310 QLSMWKYNCKNCFVLGFCOSQNEVKPGSC--PECOSAGPFVFNVEETIYQYQIRI 367
DQ 310 QLSMWKYNCKNCFVLGFCOSQNEVKPGSC--PECOSAGPFVFNVEETIYQYQIRI 367
QY 150 ELVSGTFLCLDCQTVIKD-VEQQFKYTPQNTCRPNVCANRRKRFLLDNKSRFVDFQVRI 208
DQ 150 ELVSGTFLCLDCQTVIKD-VEQQFKYTPQNTCRPNVCANRRKRFLLDNKSRFVDFQVRI 208
QY 368 QESGPKVAARLPRSKDAILADLVDSNAGDEITLGIY-----HNNYDG 413
DQ 368 QESGPKVAARLPRSKDAILADLVDSNAGDEITLGIY-----HNNYDG 413
QY 209 QETQNELRGISIPRSLEVLIRAEAVESAQAQDRCDFTGALLVDPVSKLSTPGARAETNS 268
DQ 209 QETQNELRGISIPRSLEVLIRAEAVESAQAQDRCDFTGALLVDPVSKLSTPGARAETNS 268
QY 414 SLNTANGPPV-----FATVILANHVAKDNKVAVGELTDEDV----- 450
DQ 414 SLNTANGPPV-----FATVILANHVAKDNKVAVGELTDEDV----- 450
QY 269 RVSGADGYETEGIRGLRALGVRLDSYRLVFLACHVAPTNPRFGKELRDEQTAESIKNQ 328
DQ 269 RVSGADGYETEGIRGLRALGVRLDSYRLVFLACHVAPTNPRFGKELRDEQTAESIKNQ 328
QY 451 -----KMTSLSDQDOQGEKIFASIPSIYGHEDIKRGPALAFGEPEKPGKHKVRG 504
DQ 451 -----KMTSLSDQDOQGEKIFASIPSIYGHEDIKRGPALAFGEPEKPGKHKVRG 504
QY 329 MTVWEKVFEMSDQKLNHLNCLSLPFTIHGNDEVKRGVLLMLFGGVPKRTTGGSTLSRG 388
DQ 329 MTVWEKVFEMSDQKLNHLNCLSLPFTIHGNDEVKRGVLLMLFGGVPKRTTGGSTLSRG 388
QY 505 DINVLCDGPTAKSQFLKYTEKVSSRAIFTTGGCASAVATVAVQRHPVSRWETLEAGA 564
DQ 505 DINVLCDGPTAKSQFLKYTEKVSSRAIFTTGGCASAVATVAVQRHPVSRWETLEAGA 564
QY 389 DINCVIGDPTAKSQFLKHVDESPRAVYTSGRASSAAGLTAAVVRDEESHEFVEIAGA 448
DQ 389 DINCVIGDPTAKSQFLKHVDESPRAVYTSGRASSAAGLTAAVVRDEESHEFVEIAGA 448
QY 555 LVLADRGVCLLIDFKMNDQRTSIHEAMEQOISISKAGIVTSQARCTVIAANPIGG 624
DQ 555 LVLADRGVCLLIDFKMNDQRTSIHEAMEQOISISKAGIVTSQARCTVIAANPIGG 624
QY 449 LMLADNGVCCIDEFDKMDRQDAIHEAMEQOISITKAGVKATLNARTSILAAANPVSG 508
DQ 449 LMLADNGVCCIDEFDKMDRQDAIHEAMEQOISITKAGVKATLNARTSILAAANPVSG 508
QY 625 RYDPSLTFSENVDLTEPIISRFDILCVVRDVPQDEMLARFVVGSHVRHHPNSKEEG 684
DQ 625 RYDPSLTFSENVDLTEPIISRFDILCVVRDVPQDEMLARFVVGSHVRHHPNSKEEG 684
QY 509 HYDRSKSLKQINLSAPIMSREFDLFFILVDECNVETDYAIRIVDLH----- 556
DQ 509 HYDRSKSLKQINLSAPIMSREFDLFFILVDECNVETDYAIRIVDLH----- 556
QY 685 LANGSAAPAMPNTYGVPEPLPQEVLYKVIYAKERVHPKLNQMDQDKVAKWYSDLRK--- 741
DQ 685 LANGSAAPAMPNTYGVPEPLPQEVLYKVIYAKERVHPKLNQMDQDKVAKWYSDLRK--- 741
QY 557 -----SRTEESIDRVSLDD-----IRRYLLFARQ-FPKISKESDFIVEQYKRLQRDQ 606
DQ 557 -----SRTEESIDRVSLDD-----IRRYLLFARQ-FPKISKESDFIVEQYKRLQRDQ 606
QY 742 ESMATGSPITVRHIESMSH--GGGPRAPHSAGLCDDRRRQH----- 781
DQ 742 ESMATGSPITVRHIESMSH--GGGPRAPHSAGLCDDRRRQH----- 781
QY 607 SGVTKSSWRITVRQLESIRLSSESMARMH-----CCDEVQPKHVKEAFLLNKSIRVETP 662
DQ 607 SGVTKSSWRITVRQLESIRLSSESMARMH-----CCDEVQPKHVKEAFLLNKSIRVETP 662
QY 782 -----GHPRDAGELHRHT-----EVQRHRSRKRTFYLYLS 811
DQ 782 -----GHPRDAGELHRHT-----EVQRHRSRKRTFYLYLS 811
QY 663 DVNLQDEEIQMETDEGGGVNGHADSPAPVNRNFNGSSDAQSTVSKPSLRLGFAEYC- 721
DQ 663 DVNLQDEEIQMETDEGGGVNGHADSPAPVNRNFNGSSDAQSTVSKPSLRLGFAEYC- 721
QY 812 FRDRNNELLFLPKQ-----LVABQVTVQRNRFCAQODTTEVPEKDLVDKARQIN- 861
DQ 812 FRDRNNELLFLPKQ-----LVABQVTVQRNRFCAQODTTEVPEKDLVDKARQIN- 861
QY 722 --RISNLVLHLRMEEEEDSALKRSELVNWYLKEIESEIDS-----EELINKKTIETK 775
DQ 722 --RISNLVLHLRMEEEEDSALKRSELVNWYLKEIESEIDS-----EELINKKTIETK 775
QY 862 -IHNLAFYDSELFRMKN 878
DQ 862 -IHNLAFYDSELFRMKN 878
QY 776 VVHRLT-HYDHVLIELTQ 792
DQ 776 VVHRLT-HYDHVLIELTQ 792

MCM6_HUMAN
ID MCM6_HUMAN STANDARD; PRT; 821 AA.
AC Q14566; Q13504; Q99859;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA replication licensing factor MCM6 (P105MCM).
GN MCM6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97431351; PubMed=9286856;
RA Tsuruga H., Yabuta N., Hosoya S., Tamura K., Endo Y., Nojima H.;
RT "HSMCM6: a new member of the human MCM/p1 family encodes a protein
homologous to fission yeast Mis5";
RL Genes Cells 2:381-399(1997).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98184833; PubMed=9516426;
RA Holthoff H.P., Baack M., Richter A., Ritzi M., Knippers R.;
RT "Human protein MCM6 on Hela cell chromatin";
RL J. Biol. Chem. 273:7320-7325(1998).
[3]
RP SEQUENCE OF 640-821 FROM N.A.
RX MEDLINE=97131582; PubMed=8977093;
RA Harvey C.B., Wang Y., Darmoul D., Phillips A., Mantel N.,
RA Swallow D.M.;
RT "Characterisation of a human homologue of a yeast cell division cycle
gene, MCM6, located adjacent to the 5' end of the lactase gene on
chromosome 2q21";
RL FEBS Lett. 398:135-135(1996).
CC -1- FUNCTION: MAY BE INVOLVED IN THE CONTROL OF A SINGLE ROUND OF DNA
REPLICATION DURING S PHASE. BINDS TO CHROMATIN DURING G1 AND
DETACH FROM IT DURING S PHASE AS IF IT LICENSES THE CHROMATIN TO
REPLICATE.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: BELONGS TO THE MCM FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; D84557; BAA12699.1; -;
DR EMBL; U46838; AAC50766.1; -;
DR EMBL; U67284; AAB48165.1; -;
DR EMBL; U67281; AAB48165.1; JOINED.
DR EMBL; U67282; AAB48165.1; JOINED.
DR EMBL; U67283; AAB48165.1; JOINED.
DR Genew; HGNC:6949; MCM6.
DR GK; Q14566; -;
DR MIM; 601806; -;
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0005524; F:ATP binding activity; NAS.
DR GO; GO:0006260; P:DNA replication; NAS.
DR InterPro; IPR001208; MCM.
DR Pfam; PF00493; MCM; 1.
DR ProDom; PD001041; MCM; 1.
DR SMART; SM00350; MCM; 1.
DR PROSITE; PS00847; MCM_1; 1.
DR PROSITE; PS50051; MCM_2; 1.
KW DNA replication; ATP-binding; Nuclear protein;
FT DNA replication; ATP-binding; Cell cycle; Polymorphism.
FT DOMAIN 346 553 MCM.
FT NP_BIND 396 403 ATP (POTENTIAL).
ET NP_BIND 396 403 ATP (POTENTIAL).
ET VARIANT 35 35 E -> V (IN dbSNP:3087355).
ET /FTid=VAR_014816.

FT	CONFLICT	377	387	PKTGTGTSLR -> SFDNRDRDLS (IN REF. 2).
FT	CONFLICT	495	495	A -> T (IN REF. 2).
FT	CONFLICT	738	738	MISSING (IN REF. 3).
FT	CONFLICT	790	790	L -> P (IN REF. 2).
SEQ	SEQUENCE	821 AA;	92889 MW;	F94968EB25A3E501 CRC64;

Query Match		18.0%;	Score 835;	DB 1;	Length 821;
Best local Similarity		27.5%;	Pred. No. 5.4e-45;		
Matches 236;		Conservative 155;	Mismatches 284;	Indels 182;	Gaps 22;

QY	137	LPASAAAGCAGTDEGEDEQMIESTENLEDLKGHSVREWWSMAGPRLEIHRHFKFLRTHV	196
DB	3	LAAAEFGAGSHLEVRDEVAECQKL-----	FLDFLEEF 38
QY	197	DSGHNVFKERIISDMCKENRESLVVNYEDLAAAREHVLAYFLPEA-----	PAELLQIFD 249
DB	39	SSDGEIKYLQLAELIRPERNTLVVSPVDLEQFNQOLSTTIOEEFYRVVYPLCALRALTFF	98
QY	250	EAALEVVLAMYPKYDRITNIHIVRISHLPLVEELRSLRQLHNLITSGVVTSGTGVLP	309
DB	99	KDRKEIPLA-----KDFYVAFQDLPTRHKIRELTSRIGLLTRISGGVWRTHPVHP	149
QY	310	QLSMVKYNCNKNFVLGPFQCSQNOEYKPGSC--PECQSAGPFVNMEETIYQYQIRI	367
DB	150	ELVSGTFLCLDCQTVIRD-VEQOFKYTPQNICRNPVCANRRFLDNTKSRFVDFQKRI	208
QY	368	QESPGKVAARRLPRSKDAILLADLVDSNAGDELTGTIY-----	HNNYDG 413
DB	209	QETQAEIPRCSIPRSLVILRAEVESAQAGDKCDFGTGLIVPDVSKLSTPGARAETNS	268
QY	414	SLNTANGFPV-----FATVILANHVAKKONKVAVGELTDEDV-----	450
DB	269	RVSQVDGYETEGINGRLAALGVRDLSYRLVFLACCVAPTNPFRFGKELRDEBQTAESIKNQ	328
QY	451	-----KMTISLKDQOIGKIPASIAPSIYGHEDIKRGPAALAFGEPKPNKGKKVKG	504
DB	329	MTWKEEKVPEMSQDNLYNLNLTSLPTTHGNDEVRGVLMLFGGVPKPTGEGTSLRG	388
QY	505	DINVLICDGPCTAKSQFLKVIKYSRAITFGOGASAVATVYVQRHPVSRWETLEAGA	564
DB	389	DINVICVDGSTAKSQFLKHVEEFPRAVYTSKGASSAGLTAAVVRDEESHEFVEAGA	448
QY	565	LVLADRGVCLIDFEDKMNDOQRTSIHAMEEQOISISKAGIVTSLQARCTVIAAANPTGG	624
DB	449	LMLADNGVCCTIDFEDKMDVRDQVAIHEAMEQOTISITKAGVKATLNARTSLAAANPTSG	508
QY	625	RYDPSLTFSENVDLTEDIISRFDILCVVRTVDVQDEMLARFVVGSHVRHHPSNKEEG	684
DB	509	HYDRSKLKNINISAPIMSRDFFLFDVLDECNVTDTYAIARRVLDLH-----	556
QY	685	LANGSAEAPAMPNTYGYEPLPQEVLYKVIYAKERVHPKLNOMDQDKVAKMYSDLRK---	741
DB	557	-----SRLEESIDRVSLDD-----IRYLLFARQ-FPKISKESIEDFIVEQYKHLQRDQ	606
QY	742	ESMATGSIPIVTRHIESMSH--GGGPAHPSAGLCDRRRRQH-----	781
DB	607	SGVTKSMRITVRQLESNIRLSEAMRMH---CCDEVPQKHVKEAFELLNKSIIRVETP	662
QY	782	-----GH---PRDAGELHRHTEVQRHRSMKRTFAR--YLSFPR	814
DB	663	DVNLQDDEEIQMEYDEGAGGINGHADSPAPVNGINGYNEDINQESAPKASLRLGFSYCR	722
QY	815	DNNELLIFI-----LKLQVABQVTVQRNRFGAOQDITIEVPEKDLVDKARQIN--	861
DB	723	ISNLIVLHLKRVBEEDDESALKR--SELVNNYLKEIESEIDS-----PEELINKRKITEK	776
QY	862	IHNLSAFYDSELFRMKN	878
DB	777	IHLRT-HYDHLVLIETQ	792

RESULT 15
MCM5 XENIA

ID	MCM5_XENLA	STANDARD;	PRT;	735 AA.
AC	P55862;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	DNA replication licensing factor MCM5 (CDC46 homolog) (X.CDC46).			
GN	MCM5 OR CDC46.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
LN	[1]			
LN	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RX	MEDLINE=97357318; PubMed=9214647;			
RT	Kubota Y., Mimura S., Nishimoto S., Masuda T., Nojima H., Takisawa H.;			
RT	"Licensing of DNA replication by a multi-protein complex of MCM/PI			
RT	proteins in Xenopus eggs.";			
RL	EMBO J. 16:3320-3331(1997).			
LN	[2]			
LN	SEQUENCE OF 20-735 FROM N.A.			
RC	TISSUE=Oocyte;			
RX	MEDLINE=97074651; PubMed=8917078;			
RA	Miyake S., Saito I., Kobayashi H., Yamashita S.;			
RT	"Identification of two Xenopus laevis genes, xMCM2 and xCDC46, with			
RT	sequence homology to MCM genes involved in DNA replication.";			
RL	Gene 175:71-75(1996).			
CC	-1- FUNCTION: INVOLVED IN EARLY STEP OF CHROMOSOME REPLICATION.			
CC	-1- SUBUNIT: ASSOCIATED AND INTERACTS WITH MCM2.			
CC	-1- SUBCELLULAR LOCATION: Nuclear.			
CC	-1- SIMILARITY: BELONGS TO THE MCM FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U44048; AAC60224.1; -;			
DR	EMBL; D63920; BAA09949.1; -;			
DR	PIR; T47224; PC4225.			
DR	InterPro; IPR003593; AAA_ATPase.			
DR	InterPro; IPR001208; MCM.			
DR	Pfam; PF00493; MCM; 1.			
DR	ProDom; PD001041; MCM; 1.			
DR	SMART; SM00382; AAA; 1.			
DR	SMART; SM00350; MCM; 1.			
DR	PROSITE; PS00847; MCM.1; 1.			
DR	PROSITE; PS00051; MCM.2; 1.			
KW	Transcription regulation; DNA-binding; Nuclear protein;			
KW	DNA replication; ATP-binding.			
KW	DOMAIN 332 538 MCM.			
FT	NP_BIND 382 389 ATP (POTENTIAL).			
FT	SEQUENCE 735 AA; 82435 MW; 702BA90C2F510720 CRC64;			
QY	Query Match 17.88; Score 825.5; DB 1; Length 735;			
Db	Best local Similarity 30.98; Pred. No. 1.8e-44;			
Db	Matches 231; Conservative 127; Mismatches 302; Indels 87; Gaps 18;			
QY	183 EIIHRKFNELTHVDSHGHNVEKERISDMCKENRES---LVNVEDLAAREHVLAYFLP 238			
Db	31 QLKRRFREFURQVRIQTRDTRGTFKRYDELRKRVNLGEYIEWEMEDLASFDELDADLY 90			
QY	239 EAPAEILQIFDEAALEY---VLAMPKYDRITHIHVRISHLPLVEELSLRQLHNL 295			
Db	91 KQTEHLQLLEEAQEAQVADVTRPRPAGEETIQEIQVWLRSNDANPANIRLSKSSQMSHLV 150			
QY	296 RTSGVVTSCTGVLPQLSMWKYNCKNFVLGPFQSQNQEVKPG-----SC----- 341			
Db	151 KIPGIIIAATAVRAKATKISTQCRSCNTIG-----NIAVRPGLGYAMPKRNTEQAG 204			

RESULT 15
MCM5_XENLA

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2003, 13:10:05 ; Search time 54 Seconds
(without alignments)
4262.647 Million cell updates/sec

Title: CAA47749

Perfect score: 4630

Sequence: 1 MASSPAQRRRRNDPLTSSPG.....LFRMKNFSDLRKRMILQOF 892

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_protein:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3955	85.4	904	11 Q8C2R0	Q8C2R0 mus musculus
2	3942	85.1	904	11 Q08971	Q08971 mus musculus
3	3634.5	78.5	811	4 Q8N2V1	Q8N2V1 homo sapien
4	3551.5	76.7	886	13 Q42588	Q42588 xenopus lae
5	3192.5	69.0	880	13 Q8JGR5	Q8JGR5 brachydanio
6	2254	48.7	881	5 Q9XXI9	Q9XXI9 caenorhabdi
7	2132	46.0	871	5 Q9B122	Q9B122 dugesia jap
8	1945.5	42.0	936	10 Q9LPD9	Q9LPD9 arabidopsis
9	1887.5	40.8	865	10 Q8H0G9	Q8H0G9 nicotiana t
10	1864.5	40.3	889	3 Q42723	Q42723 emericella
11	1604.5	34.7	365	4 Q9BWF4	Q9BWF4 homo sapien
12	1525	32.9	780	5 Q8SS42	Q8SS42 encephalito
13	1277	27.6	883	5 Q9U446	Q9U446 entamoeba h
14	1267	27.4	971	5 Q8ILU7	Q8ILU7 plasmodium
15	1262	27.3	971	5 Q9GR05	Q9GR05 plasmodium
16	1253	27.1	491	10 Q96275	Q96275 arabidopsis

17	913	19.7	686	17 Q9UXG1	Q9uxg1 sulfolobus
18	909.5	19.6	680	17 Q82V88	Q82v88 pyrobaculum
19	909.5	19.6	699	17 Q8YFER1	Q8yfer1 aeropyrum p
20	872	18.8	863	4 Q8NEH1	Q8nehl homo sapien
21	858	18.5	720	10 Q9SIV8	Q9siv8 arabidopsis
22	856	18.5	849	10 Q8RYQ2	Q8ryq2 oryza sativ
23	855	18.5	776	10 Q65400	Q65400 arabidopsis
24	853	18.4	776	10 Q9FL33	Q9fl33 arabidopsis
25	853	18.4	862	11 Q9D077	Q9d077 mus musculu
26	852	18.4	862	11 Q8C1Z0	Q8c1z0 mus musculu
27	846	18.3	862	11 Q921D5	Q921d5 mus musculu
28	843.5	18.2	858	13 Q42589	Q42589 xenopus lae
29	839.5	18.1	768	10 Q9SX03	Q9sx03 zea mays (m
30	839.5	18.1	768	10 Q9SX04	Q9sx04 zea mays (m
31	825.5	17.8	701	17 Q8TSM4	Q8tsm4 methanosarc
32	823.5	17.8	666	17 Q27798	Q27798 methanobact
33	821.5	17.7	824	13 Q73710	Q73710 xenopus lae
34	819	17.7	819	5 Q9XYU1	Q9xyul drosophila
35	812	17.5	819	5 P91675	P91675 drosophila
36	810.5	17.5	701	17 Q8PVX1	Q8pvx1 methanosarc
37	807.5	17.4	796	13 Q42590	Q42590 xenopus lae
38	800	17.3	733	5 Q9VGW6	Q9vgw6 drosophila
39	797	17.2	823	5 Q95XQ8	Q95xq8 caenorhabdi
40	796.5	17.2	727	10 Q80786	Q80786 arabidopsis
41	788	17.0	720	13 Q42592	Q42592 xenopus lae
42	788	17.0	726	5 Q8SRX5	Q8srx5 encephalito
43	786	17.0	710	10 Q8H1A2	Q8hia2 pisum sativ
44	784	16.9	586	17 Q29733	Q29733 archaeoglob
45	784	16.9	812	5 Q9XVR7	Q9xvr7 caenorhabdi

ALIGNMENTS

RESULT 1

Q8C2R0

ID Q8C2R0 PRELIMINARY; PRT; 904 AA.

AC Q8C2R0;

DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Mini chromosome maintenance deficient 2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NOD; TISSUE=Thymus;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium.

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RL Nature 420:563-573(2002)."

DR ENBL: AK088156; BAC40178.1; -

SQ SEQUENCE 904 AA; 102077 MW; 7B61C13DADICAC58 CRC64;

Query Match

Best Local Similarity 85.4%; Score 3955; DB 11; Length 904;

Matches 788; Conservative 23; Mismatches 61; Indels 42; Gaps 7;

QY 2 ASSPA-QRRRGNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGEGPLEEEDG 60

||||| ||||| :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

10 ASSPAQRRRISDPLTSSPGRRRADALTSSPGRDLPPEDESEGLLTGEGPMEEEDG 69

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 EELIGCMERYRAIPELDVAEAGLADDEDEVELTASRREA-----DGPCGNVTG 113

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 70 EELIGCMERYRPIPELDVYAEGLADDEDEVELTASQREAAERTWRDRAGR--- 126

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 114 SWPGLGACAVGSCMTAMRTRTSALPASAAAG-----AGTDEGEDEQMIESIE 161

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 127 ---GLG-----RMRRGLLYDSSEDEERPAKRKRHRVERATEDGEDEQMIESIE 172

```
QY 162 NLEDLKGHSVREWSMAGPRLEIHHRFKNFRLTHVDSHGHNVPKERSIDMCKENRESLIV 221
DB 173 NLEDLKGHSVREWSMAGPRLEIHHRFKNFRLTHVDSHGHNVPKERSIDMCKENRESLIV 232
QY 222 NYEDLAAREHVLAVFLPAPAEQLQIFDEAALEVVLVAMYPKYDRITNIHVRISHLPLVE 281
DB 233 NYEDLAAREHVLAVFLPAPAEQLQIFDEAALEVVLVAMYPKYDRITNIHVRISHLPLVE 292
QY 282 ELRSRLQLHLNQLRTSGVTSCTGVLPQLSNVKNCKNFVLGPFQCSQNEQVKPGSC 341
DB 293 ELRSRLQLHLNQLRTSGVTSCTGVLPQLSNVKNCKNFVLGPFQCSQNEQVKPGSC 352
QY 342 PECOSAGPEYFVMEETIYQYORIQESPGKVAARLPRSKDAILLADLVDSNAGDEI 401
DB 353 PECOSAGPEYFVMEETIYQYORIQESPGKVAARLPRSKDAILLADLVDSNAGDEI 412
QY 402 ELTGIIYHNNYDGLNTANGFPVFATVILANHVAKDNKVAVGELTDEVDKMTSLSKDQ 461
DB 413 ELTGIIYHNNYDGLNTANGFPVFATVILANHVAKDNKVAVGELTDEVDKMTSLSKDQ 472
QY 462 IGEKIFASIASIYGHEDIKRGPALALFGGPKNPGGKHVYRGDINVLCCGDPGTAKSQF 521
DB 473 IGEKIFASIASIYGHEDIKRGPALALFGGPKNPGGKHVYRGDINVLCCGDPGTAKSQF 532
QY 522 LKYTEKVSSRAIFTTGGASAVATYVYORHPVPSREWTLFAGALVADRGVCLIDEFDKM 581
DB 533 LKYTEKVSSRAIFTTGGASAVATYVYORHPVPSREWTLFAGALVADRGVCLIDEFDKM 592
QY 582 NDQDRTSIEHAMEQOSISISKAGIVTSLOARCTVIAAANPIGGRYDPSLTSENVDLTEP 641
DB 593 NDQDRTSIEHAMEQOSISISKAGIVTSLOARCTVIAAANPIGGRYDPSLTSENVDLTEP 652
QY 642 IISRFDIICVYRVDPVQDEMLARFVYVSHVRHPSNKEEGLANGSAAPAMPNTYGV 701
DB 653 IISRFDIICVYRVDPVQDEMLARFVYVSHVRHPSNKEEGLANGSAAPAMPNTYGV 712
QY 702 EPLPQEVLLKYYIAKERVHPLNOMDQDKVAKYMSDLRKESMATGSPITVRHIESMSH 761
DB 713 EPLPQEVLLKYYIAKERVHPLNOMDQDKVAKYMSDLRKESMATGSPITVRHIESMSH 772
QY 762 GGGPRAHPSAGLCDDRRRQHGHPDAGELHRTVEVQRH---RSMKTFARVLSRRDNE 818
DB 773 MA--EAAHMLHLDVYMEDDYMNAIRVMMESFDTQKFSVMSRMTFARYLSRRDND 830
QY 819 LLLFLKOLVAEQVYQNRNFGAQQDTIEVPEKDLVDKARQINHNLSAFYDSELFMRNK 878
DB 831 LLLFLKOLVAEQVYQNRNFGAQQDTIEVPEKDLVDKARQINHNLSAFYDSDLFMRNK 890
QY 879 FSHDLKRKMILOQF 892
DB 891 FSRDLKRKLILQOF 904
```

RESULT 2

```
008971
ID 008971 PRELIMINARY; PRT; 904 AA.
AC 008971;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE BM28 homolog.
GN CMMD2 OR CDCL1 HOMOLOG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=98037564; PubMed=9371513;
RA Spanjaard R.A., Lee P.J., Sarkar S., Goedegebuure P.S., Eberlein T.J.;
RT "Clone 10d/BM28 (CDCL1), an early S-phase protein, is an important
```

```
growth regulator of melanoma.";
RT Cancer Res. 57:5122-5128(1997).
DR EMBL; AF004105; AAC16250.1; -.
DR MGD; MGI:105380; Mcmd2.
DR InterPro; IPR001208; MCM.
DR Pfam; PF00493; MCM; 1.
DR ProDom; PD001041; MCM; 1.
DR SMART; SM00350; MCM; 1.
DR PROSITE; PS00847; MCM_1; 1.
DR PROSITE; PS00551; MCM_2; 1.
SQ SEQUENCE 904 AA; 101963 MW; 8B982C67FABBF0D0 CRC64;

Query Match 85.1%; Score 3942; DB 11; Length 904;
Best Local Similarity 86.7%; Pred. No. 2.7e-275;
Matches 784; Conservative 25; Mismatches 73; Indels 22; Gaps 17;

QY 2 ASSPA-QRRRGNDPLTSSPGSSRRRTDALTSPPGRDLPPFEDESGLLGTGEPLEEDG 60
DB 10 ASSPARQRRRISDPLTSSPGSSRRRADALTSPGRDLPPFEDESGLLGTGEPMEEDG 69
QY 61 EELIGDGMERYRAIPELDAYEAGLALDDEDEVELTASRREA-----DGCQGVTVG 113
DB 70 GELIGDGMERYRPIPELDVYEAGLALDDEDEVELTASQREAAERTMRQDREAGR--- 126
QY 114 SWPGLGACAVGSCMTAMRTRSAIPASAAGA--GTEDGEDEDEQMIENLEDLKGHSV 171
DB 127 ---GLGRMRRLPYDSSSEDEER-PARKRRHVRATEDGEDEDEMIENLEDLKGHSV 182
QY 172 REWMSMAGPRLEIHHRFKNFRLTHVDSHGHNVPKERSIDMCKENRESLVVYEDLAAREH 231
DB 183 REWMSMAGPRLEIHHRFKNFRLTHVDSHGHNVPKERSIDMCKENRESLVVYEDLAAREH 242
QY 232 VLAYFLPAPAEQLQIFDEAALEVVLVAMYPKYDRITNIHVRISHLPLVEELRSRLQLHL 291
DB 243 VLAYFLPAPAEQLQIFDEAALEVVLVAMYPKYDRITNIHVRISHLPLVEELRSRLQLHL 302
QY 292 NOLIRTSVVTSCITGVLPQLSNVKNCKNFVLGPFQCSQNEQVKPGSCPCQAGPFE 351
DB 303 NOLIRTSVVTSCITGVLPQLSNVKNCKNFVLGPFQCSQNEQVKPGSCPCQAGPFE 362
QY 352 VNMETIYQYORIQESPGKVAARLPRSKDAILLADLVDSNAGDEIELTGYIHNNY 411
DB 363 INMETIYQYORIQESPGKVAARLPRSKDAILLADLVDSNAGDEIELTGYIHNNY 422
QY 412 DGSNTANGFPVFATVILANHVAKDNKVAVGELTDEVDKMTSLSKDQIQEIKFASIA 471
DB 423 DGSNTANGFPVFATVILANHVAKDNKVAVGELTDEVDKMTSLSKDQIQEIKFASIA 482
QY 472 PSYIGHEDIKRGPALALFGGPKNPGGKHVYRGDINVLCCGDPGTAKSQFLYIEKVSRR 531
DB 483 PSYIGHEDIKRGPALALFGGPKNPGGKHVYRGDINVLCCGDPGTAKSQFLYIEKVSRR 542
QY 532 AIFTTGGASAVATYVYORHPVPSREWTLFAGALVADRGVCLIDEFKMNDQDRTSIEH 591
DB 543 AIFTTGGASAVATYVYORHPVPSREWTLFAGALVADRGVCLIDEFKMNDQDRTSIEH 602
QY 592 AMEQOSISISKAGIVTSLOARCTVIAAANPIGGRYDPSLTSENVDLTEPIISRFDILCV 651
DB 603 AMEQOSISISKAGIVTSLOARCTVIAAANPIGGRYDPSLTSENVDLTEPIISRFDILCV 662
QY 652 VRDTPVQDEMLARFVYVSHVRHPSNKEEGLANGSAAPAMPNTYGVPEPLPQEVLLK 711
DB 663 VRDTPVQDEMLARFVYVSHVRHPSNKEEGLANGSAAPAMPNTYGVPEPLPQEVLLK 722
QY 712 YIIYAKERVHPLNOMDQDKVAKYMSDLRKESMATGSPITVRHIESMSHGGPRAHPSA 771
DB 723 YIIYAKERVHPLNOMDQDKVAKYMSDLRKESMATGSPITVRHIESMSHMA--BAHARM 780
QY 772 GLCDRRRRQHGHPDAGELHRTVEVQRH---RSMKTFARYLSRRDNEALLFLKQLV 828
DB 781 HLRDYVMDYNNMNAIRVMMESFDTQKFSVMSRMTFARYLSRRDNEALLFLKQLV 840
QY 829 AEQVYQNRNFGAQQDTIEVPEKDLVDKARQINHNLSAFYDSELFMRNKFSHDLKRKMI 888
```

[illegible]

```
Db 233 LPEAPAEMLKTFDEAAKEVWLVMPKYDRIAREIHVRISHLPVLEELSLRQLNLQIR 292
QY TSGVVTCTGVLPLQSLVMKYKNCNCFVLGPFQSONOEKVPKPCQSCAGPPEVNME 356
Db 293 TSGVVTCTGVLPLQSLVMKYKNCNCFVLGPFQSONOEKVPKPCQSCAGPPEVNME 352
QY TIYQNYQIRIQESPGKVAARLRPSKDAILLADLVDSCKPGDELTGTGIYHNYDGLN 416
Db 353 TVYQNYQIRITIOESPGKVAARLRPSKDAILLADLVDSCKPGDELTGTGIYHNYDGLN 412
QY TANGFPVPATVILANHVAKKONKVAAGBELTDEDVKMTSLSKDOQIGKIFASTAPSIY 476
Db 413 TANGFPVPATVILANHVAKKONKVAAGBELTDEDVKMTSLSKDOQIGKIFASTAPSIY 472
QY HEDIKRGPALALFGEKPNKPGKHVRGDIINVLGCGDPTAKSOFKLYIEKVSSRAFTT 536
Db 473 HEDIKRGPALALFGEKPNKPGKHVRGDIINVLGCGDPTAKSOFKLYIEKVSSRAFTT 532
QY GOGASAVATYVORHPVSRWTLFAGALVLDAGVCLIDFEDKMDQDRTSIHEAMEQQ 596
Db 533 GOGASAVATYVORHPVSRWTLFAGALVLDAGVCLIDFEDKMDQDRTSIHEAMEQQ 592
QY SISISKAGIVTSLQARCTVIAAANPIGGRYDPSLTFSENVDLTPEIISRFDILCVVRTV 656
Db 593 SISISKAGIVTSLQARCTVIAAANPIGGRYDPSLTFSENVDLTPEIISRFDILCVVRTV 652
QY DPVQDEMLARFVGVSHVRRHPSNKEEGLANGSAEAMPNTYGVPELPQEVKLYIYA 716
Db 653 DPVQDEMLARFVGVSHVRRHPSNKEEGLANGSAEAMPNTYGVPELPQEVKLYIYA 709
QY KERVHPKLNQMDQKVAKMYSDLRKESMATGSIPTVVRHIESMHSBGGRAPHSAGLCDR 776
Db 710 KERVHPKLNQMDQKVAKMYSDLRKESMATGSIPTVVRHIESMHSBGGRAPHSAGLCDR 767
QY RRRQHGPRDAGELHRHTEVQRH--RSMRKTARYLSFRDNNEELLFLKQLVAEQVT 833
Db 768 RRRQHGPRDAGELHRHTEVQRH--RSMRKTARYLSFRDNNEELLFLKQLVAEQVT 827
QY YORNRFGAQQDTIEVPEKDLVDKARQINIHLSAFYDSELPKMKFSDHDKRMILQOF 892
Db 828 YORNRFGAQQDTIEVPEKDLVDKARQINIHLSAFYDSELPKMKFSDHDKRMILQOF 886

RESULT 5
Q8JGR5 PRELIMINARY; PRT; 880 AA.
AC Q8JGR5;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DE DNA replication licensing factor.
GN MCM2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22035902; PubMed=12006978;
RA Golling G., Amsterdam A., Sun Z., Antonelli M., Maldonado E., Chen W.,
RA Burgess S., Haldi M., Artzt K., Farrington S., Lin S.-Y., Nissen R.M.,
RA Hopkins N.;
RT "Insertional mutagenesis in zebrafish rapidly identifies genes
RT essential for early vertebrate development.";
RL Nat. Genet. 31:135-140(2002).
DR EMBL; AY099531; AAM28219.1;
DR InterPro; IPR001208; MCM.
DR Pfam; PF00493; MCM; 1.
DR ProDom; PD01041; MCM; 1.
DR PROSITE; PS00847; MCM_1; 1.
DR PROSITE; PS00851; MCM_2; 1.
```

```
SQ SEQUENCE 880 AA; 98304 MW; 34C496C2ED14ABBC CRC64;
Query Match
Best Local Similarity 69.0%; Score 3192.5; DB 13; Length 880;
Matches 658; Conservative 64; Mismatches 130; Indels 45; Gaps 16;
QY 15 LTSPGSSRRRTDALTSPPGRDLPPFEDESEGLLGTGTEGLEEEDEBELIGDMEROYRA 74
Db 10 MATSPTRGSRGD-LTSPGRDLPPFEDESEGLLGLDTPDEDDDDGEEELIGDAMERDVR 68
QY 75 IPELDAYEAGLALDDEDEVELTASRRRAADPCGTGTSWPGLCACAVGSCM----- 127
Db 69 VPEDRYEAGLD-EDELSSELSPSARAEAAAMRR-RDRQGGLGMGRGLGLLYDSDE 126
QY 128 TAMRRT-RSALPASAGAGTDEDEQMTESLENEDLKGHVSREWVSMAGPRLETHH 186
Db 127 DDKRPTKRQVLAERAAEGGAMEG-EDEEMTESLENL---GHE-----GPHGARVGYG 176
QY 187 R-----FKNFLRTHVDSHGHNFKERISDMCKENRESLVVNYEDLAARHVLAYFLP 238
Db 177 RTTGTDLPLKFNFLRTHVDEHGHNFKERISDMCKENKESLLVNYEDMLPRACV-GVEST 235
QY 239 EAPAELOIFDEAALEVVLAMYPKYDRITNIHVRISHLPVLEELSLRQLHLNQLITS 298
Db 236 KAPAEMLKIFDEAAKEVVLAMYPKYDRIAEIHVRIGNLP---LGSRLQ---QLIPTS 288
QY 299 GWVTCTGVLPLQSLVMKYKNCNCFVLGPFQSONOEKVPKPCQSCAGPPEVNMEETI 358
Db 289 GWVTCTGVLPLQSLVMKYKNCNCFVLGPFQSONOEKVPKPCQSCAGPPEVNMEETV 348
QY 359 YQNYQIRIQESPGKVAARLRPSKDAILLADLVDSCKPGDELTGTGIYHNYDGLN 418
Db 349 YQNYQIRIQESPGKVAARLRPSKDAILLADLVDSCKPGDELTGTGIYHNYDGLN 408
QY 419 NGFPVFATVILANHVAKKONKVAAGBELTDEDVKMTSLSKDOQIGKIFASTAPSIY 478
Db 409 NGFPVFATVILANHVAKKONKVAAGBELTDEDVKMTSLSKDOQIGKIFASTAPSIY 468
QY 479 DIKRGPALALFGEKPNKPGKHVRGDIINVLGCGDPTAKSOFKLYIEKVSSRAFTTQ 538
Db 469 DIKRGPALALFGEKPNKPGKHVRGDIINVLGCGDPTAKSOFKLYIEKVSSRAFTTQ 528
QY 539 GASAVATYVORHPVSRWTLFAGALVLDAGVCLIDFEDKMDQDRTSIHEAMEQQSI 598
Db 529 GASAVATYVORHPVSRWTLFAGALVLDAGVCLIDFEDKMDQDRTSIHEAMEQQSI 588
QY 599 SISKAGIVTSLQARCTVIAAANPIGGRYDPSLTFSENVDLTPEIISRFDILCVVRTV 658
Db 589 SISKAGIVTSLQARCTVIAAANPIGGRYDPSLTFSENVDLTPEIISRFDILCVVRTV 648
QY 659 VQDEMLARFVGVSHVRRHPSNKEEGLANGSAEAMPNTYGVPELPQEVKLYIYAKE 718
Db 649 VQDEMLARFVGVSHVRRHPSNKEEGLANGSAEAMPNTYGVPELPQEVKLYIYAKE 705
QY 719 RVHPKLNQMDQKVAKMYSDLRKESMATGSIPTVVRHIESMHSBGGRAPHSAGLCRRR 778
Db 706 RVHPKLNQMDQKVAKMYSDLRKESMATGSIPTVVRHIESMHSBGGRAPHSAGLCRRR 763
QY 779 RQHGPRDAGELHRHTEVQRH--RSMRKTARYLSFRDNNEELLFLKQLVAEQVTQ 835
Db 764 RQHGPRDAGELHRHTEVQRH--RSMRKTARYLSFRDNNEELLFLKQLVAEQVTQ 823
QY 836 RNRFCAQDQDTIEVPEKDLVDKARQINIHLSAFYDSELPKMKFSDHDKRMILQOF 892
Db 824 RNRFCAQDQDTIEVPEKDLVDKARQINIHLSAFYDSELPKMKFSDHDKRMILQOF 880

RESULT 6
Q9XXI9 PRELIMINARY; PRT; 881 AA.
ID Q9XXI9
AC Q9XXI9;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
```


Db 197 YSERIGTMAQANKSLTINYBHLVAVEQALAYFLPEAPOGMLEIFDEVAFAVTLSRPYNY 256
Qy 264 DRITNHHVTRISHPLVEELRSRLQHLNQLIRTSVGVVTCGTGLV---POLSMWKYCNKC 320
Db 257 KTITESHVTRISHPLVEELRSRLQHLNQLIRTPEL---CQVLLFCPLNLVQFNCR 313
Qy 321 CNFVLGPFQSQN--QEVKPGSCPCQAGPEPEVMEETIYQYQRIQESPGKVAARRL 379
Db 314 CGCTIGPTQTDSTNEIKPSTPCDQCQSPPEINVEKTYKNYQRIQESPGVPAGRL 373
Qy 380 PRSKDAILLADLVSCNAGDEIELTGIIYHNNYDGLTANGFPFPAIVLANHHVAKDNK 439
Db 374 PRSKDAILLADLVSCNAGDEIDITGIIYHNNYDGLTANGFPFPAIVLANHHVAKDNK 433
Qy 440 VAGELTDEDVKMTISLTKDOQIGEKIFASITAPSIYGHEDIKRGPALALFGGPKNPGK 499
Db 434 LILSGVTDDEDIVNTONLAKDERLDRILRSITAPSIYGHENIKRAIALSLFGGVAKTKGQK 493
Qy 500 HKVRGDIVNLGCGDPTAKSOFKLYIEKVSSRAITFTTGGQASAVATYVORHPVSRWT 559
Db 494 LRGRGDIVNLGCGDPTAKSOFKLYIEKVSSRAITFTTGGQASAVATYVORHPVSRWT 553
Qy 560 LEAGALVLADRGVCLIDFEDKMDQDRTSIEHMEQOSISISKAGIVTSLOARCTVAAA 619
Db 554 LEAGALVLADRGVCLIDFEDKMDQDRTSIEHMEQOSISISKAGIVTSLOARCTVAAA 613
Qy 620 NPIGRYDPSLTFSENVDLTPITSRFDILCVVRDTPVQDDEMLARFVGVSHVRHP-- 677
Db 614 NPIAGRYDTSRFDNDVLTSPILSRFDILCVVRDTPVQDDEMLARFVGVSHVRHP-- 673
Qy 678 SNKEEGLANGSAEAPAMPNTY-----GVEPLQVEVLKYYIYAKERVHPKLNQMDQKV 732
Db 674 SEEBKEVVKRLAEGALMANTSDSEAGDLQAIPLLELLKYYIYAKERVHPKLNQMDQKV 733
Qy 733 AKMYSDLRKESMATGSPITVTHIESHSGGPRAPHSAGLCDRRRHQGHPRAGELHR 792
Db 734 CEVYEURKASKENGIDVTYRFSMT-----RUSEAH-----AKLHF 772
Qy 793 HTEVQRH-----RSMRKTARYLSFRDNNELLFILKQLVAE 830
Db 773 REVNEEDVNMAIRVLESFISTEKYSIMKMLENKRHYLNYKNDHELINFILKQFHD 832
Qy 831 QVTYORNFQAQDT-TEVPKDLVDKARQINI 862
Db 833 QLTTLRAEGKSLNLSIIEITKSLDKVRRMNI 865

RESULT 8
Q9LPD9 PRELIMINARY; PRT; 936 AA.
AC Q9LPD9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE T12C22.19 protein.
GN T12C22.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosoids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Liu S., Vaysberg M., Sakano H., Lee J., Lenz C., Pham P., Toriumi M.,
RA Yu G., Chin C., Chou J., Choi E., Chung M., Gonzalez A., Howng B.,
RA Liu A., Altai H., Brooks S., Buehler E., Chao Q., Conn L., Conway A.,
RA Hansen N., Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M.,
RA Nguyen M., Palm C., Shinn P., Southwick A., Davis R., Ecker J.,
RA Federspiel N., Theologis A.;
RT "The sequence of BAC T12C22 from Arabidopsis thaliana chromosome 1.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC020576; AAF/8275.1; -.

DR InterPro: IPR001208; MCM.
DR Pfam: PF00493; MCM; 1.
DR ProDom: PD01041; MCM; 1.
DR SMART: SM00350; MCM; 1.
DR PROSITE: PS00847; MCM_1; 1.
DR PROSITE: PS00501; MCM_2; 1.
SQ SEQUENCE 936 AA; 105583 MW; EDED5161964DBC69 CRC64;

Query Match 42.0%; Score 1945.5; DB 10; Length 936;
Best Local Similarity 45.3%; Pred. No. 2.9e-131;
Matches 430; Conservative 150; Mismatches 268; Indels 101; Gaps 20;

Qy 12 NDPLT-SSPGRSSRRTDALTSPPGRDLPPFDESEGLGT-----EGPLEEEDGEE 62
Db 8 NEPSASPSPAGNTDQLPISTSONSENSEDEEAAVDTQVIRDEPDEADEEED 67
Qy 63 LIGGMEDYRAIFELDAYEAGL--ALDDE-DVEELTASREAADGPC----- 108
Db 68 LFNDTFMNDYRKMDENQOYESNGIDDSVDDERDLQGMALDRR--AADADLADRENLANRK 126
Qy 109 -----GTVTGSWPLGACAVGSCMTAMRRTRSALPASAAAGAGTGDG-----EED 153
Db 127 LPHLLHNDSDDN-----YRPSKRSRTTVP--PRNGGDPDGNPSSPGVSPQD 174
Qy 154 EOMTIESIENLED-----LKGHSVREVMVSMAGPRLEIHHKFNLRTHV--- 196
Db 175 ISMTDQDDYQDEDDNDDEAEFEMVRIQG--TLREWMRDEVRRTAKKFKDFLLTYVVKP 233
Qy 197 DSHGHNVPKERISDMCKENRESLVVNYEDLAAREHVLAYFLPEAPAEILLQIFDEAALEVV 256
Db 234 NENGDIETVRLINEMVSANKCSLEIDYKEFTVHPNITAIWLADAPQPVLEVMEEVSEKVI 293
Qy 257 LAMYPKYDRINHTHVRISHPLVEELRSRLQHLNQLIRTSVGVVTCGTGLVLPOLSMVKY 316
Db 294 FDLHPNKNHTKIYVRVNLVNDQIRNIQIHLNTHIRIGGVVTRSGVFPQLOQVKY 353
Qy 317 NCNKNFVLGPFQSQNQEVKPGSCPCQAGPEPEVMEETIYQYQRIQESPGKVA 376
Db 354 DCNKGAVLGPFQSQNQEVKPGSCPCQAGPEPEVMEETIYQYQRIQESPGKVA 413
Qy 377 RRLPRSDAILLADLVSCNAGDELTGTIYHNNYDGLTANGFPFPAIVLANHHVAKK 436
Db 414 GRLPKHVEVILLNDLIDCARPEEIEVGTIYHNNYDGLTANGFPFPAIVLANHHVAKK 473
Qy 437 DNKVAVGELTDEDVKMTISLTKDOQIGEKIFASITAPSIYGHEDIKRGPALALFGGPKN 496
Db 474 QDLFSAYKLTQEDKTQIEELSKDPRIVERIKSIAPSIYGHEDIKTALALAMFGGOEKNI 533
Qy 497 GGHKVRGDIVNLGCGDPTAKSOFKLYIEKVSSRAITFTTGGQASAVATYVORHPVSR 556
Db 534 KGHRLRGDIVNLGCGDPTAKSOFKLYIEKVSSRAITFTTGGQASAVATYVORHPVSR 593
Qy 557 EMTLEAGALVLADRGVCLIDFEDKMDQDRTSIEHMEQOSISISKAGIVTSLOARCTV 616
Db 594 EMTLEAGALVLADRGVCLIDFEDKMDQDRTSIEHMEQOSISISKAGIVTSLOARCTV 653
Qy 617 AAANPIGRYDPSLTFSENVDLTPITSRFDILCVVRDTPVQDDEMLARFVGVSHVRHH 676
Db 654 AAANPVGGRYDSSKSFAGNVELTDPILSRFDILCVVRDTPVQDDEMLARFVGVSHVRHH 713
Qy 677 P-----SNKEEGLANGS--AAEPAMPNTYQVPELQVEVLKYYIYAKERVHPKLNQMDQ 729
Db 714 PKGKMEDSDPEDGTGSSGSDTP-----EVLQNLKLYTYSKLYVFPKLDELDA 765
Qy 730 DKVAKMYSDLRKESMATGSPITVTHIESHSGGPRAPHSAGLCDRRRHQGHPRDAGE 789
Db 766 KKLETVANLRRESMNGGVSIATRHLESMTIRMS--EAHARMHL-----ROYTVEEDVNM 818
Qy 790 -----LHRHTEVQR---HRSMRKTARYLSFRDNNELLFILKQLVAEQTYQNRREGA 841
Db 819 AIRVLLDSFISTQKFGVORTLRESKRYITTKDFNSLLVLLKELVKNALKFEEISS 878
Qy 842 QQ--DTIEVPEKDLVDKARQINIHLNSAFYDSELSFRMKNKFSHDLKRRKI 888


```
QY 94 E-----ELTASRRRRAADGPGCTVTGSMPLGACAVGSCMTAMRR-----TRSLAP 138
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
76 DDDFIDDELDYEDNLKARRAE-----RHMOMORKEGKYKKNKEW 116
QY 139 ASASACAGTDEED-----EQMIESIENLE-----DLKG-HSVREWVSMAGPRL 182
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 KTELOEGDDEEDIFDKVAKRRENHLTAETDIPDLSNLSAKICLSVNPDKDV 176
QY 183 EIHHR-----FKNFL-----RTHVDSH--GHNVFERISDMCK 213
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 VDERVQQAADTCFRYFLHKFSIKDSMGLNIDESNTELEHEEHMNSSHOYYYDKTEKMIL 236
QY 214 ENRESUVNYEDLAAREHV--LAYFLPEAPABELLQIFDEAALVVLAMYPKY--DRITNH 269
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
237 NDKHTLIVSAKHL-IQFHCENLVQWIEFKPEQILEVHECLAVEAYRISPKLYKGRICK- 294
QY 270 IHVRISHLPVBELSRLQHLNQLIRTSQVTSCTGYLPQLSMVKYCNKCNFVLGP-- 327
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
295 --VVLRDWPYQLRNLRLCTELNLIKVTGVCIKRGYVLPKLRVMYKLCNSCDTTLSEYP 352
QY 328 -FCQSONQEVKPGSCPEOSAGPFVFNMEETIYQYORIRIOESPGKVAARLPRSKDAI 386
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
353 IYFADGKKVLPRLPRCPHCOSA-TFSVDRIKTYATYDQKITLOESPCSVPAGRAPOREVV 411
QY 387 LLADLVDSNAGDEIELTGIIYHNNDGSLTANGPVPFATVILANHVAKKDNKVAVGELT 446
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
412 VTGDLVDKVKPGEEVEVLGIYKYDIGNIKYGPILQITEFANNIERKED-IQLSELT 470
QY 447 DEDVKMITSLDQOQIGEFIFASIAPSIYGHEDIKRGPALALFPGEPK----- 494
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
471 EDDIKDLKSLDPNIRIRIITSIAPAIWGHDKITSIAYALFGVQKGGKSFSSKNNET 530
QY 495 -----NPGGKHVRGDIINVLLCGDPTAKSQFLKYEKVSSRAIETTGOGASA 542
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
531 NNFQVQNRDILNFKGCHTIRGDIINVLLGDPGLGKSQVLQYVHKTNLRTVITTKGASA 590
QY 543 VAVTAYVQRHPVSRWTLBAGALVLADRGVCLIDEFKMNDQDRTSIHEAMEQQQSISISK 602
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
591 VGLTAGVRKDHTTNEWTLEGGALVLADEGICIDEFDKMTDKDRVSIHEAMEQQQSISISK 650
QY 603 AGIVTSLQARCVIAAANPIGRYDPSLTFSENVDLTFPIISRFDILCVVRTDVPVQDE 662
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
651 AGIVTTLRARCARIAAANPIGRYNPSLTFKENVDLSDPILSRFDLITVLRDIPNVDEF 710
QY 663 MLARFVVGSHVRHPS-----NKEEGLANGSAAEPAMPNTYGVPELPQEVLYKYYIY 715
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
711 YLAEYVVTNHQLSHPKLENTQYQRIENLKN-----VIVSSAYEPIQDQLLYYYIY 764
QY 716 AKERVHPKLNQMD---ODKVAKMYSDLRKESMATGSIPTVTRHIESMSHGGGPRAHPSA 771
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
765 ARTNCKPSLSDVPYABISAKLSNFYSRVQKACASGGYPLTLRHIESIIRIAEANA--- 820
QY 772 GLCDRRRROHG-HPRD-----AGELHRHTEVQRH---RSMRKTFFARYLSFRDNNELLF 822
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
821 ----KMRLSHQIYSKDVDAIAITLLESYVSCORFAVAKOLSKFEFARYALFRGGEVLRE 876
QY 823 ILKOLVAEQVTYORNFGAQQDTIEVPEKDLVDKARQIINHLSAFYDSELF----RMNK 878
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
877 LLRRTVQMIQRQNLKNASAKDFONDSSEGTSEAEILPNV--FLPLHIFMKTAMQNK 934
QY 879 FS 880
Db : :
935 FS 936
```

Search completed: September 12, 2003, 13:13:50
Job time : 59 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2003, 13:12:35 ; Search time 20 Seconds

(without alignments)
1887.064 Million cell updates/sec

Title: CAA47749

Perfect score: 4630

Sequence: 1 MASSPAQRGRNDPLTSSPG.....LFRMKFSDHDKRMILQQF 892

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCFUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	774	16.7	720	4	US-09-604-605-2
2	140	3.0	918	4	US-09-198-452A-1072
3	128.5	2.8	1515	4	US-09-328-352-6100
4	126.5	2.7	871	4	US-09-134-001C-3979
5	125	2.7	15281	2	US-08-471-1119A-2
6	121.5	2.6	516	4	US-09-252-991A-24044
7	121	2.6	875	4	US-09-107-532A-4537
8	121	2.6	1162	4	US-09-134-001C-4008
9	118.5	2.6	1589	3	US-08-755-587-189
10	117.5	2.5	1698	3	US-09-315-793-12
11	117	2.5	616	2	US-09-001-826-5
12	117	2.5	616	4	US-09-264-854-5
13	116	2.5	771	4	US-09-252-991A-17243
14	116	2.5	1151	3	US-08-840-006-6
15	116	2.5	1200	3	US-08-840-006-5
16	115	2.5	780	4	US-09-252-991A-18846
17	115	2.5	1481	4	US-09-231-899-70
18	113	2.4	1105	3	US-08-999-774A-2
19	112.5	2.4	483	4	US-09-252-991A-31250
20	111.5	2.4	499	4	US-09-328-352-5415
21	111.5	2.4	1388	4	US-09-252-991A-20237
22	110.5	2.4	469	2	US-08-989-478-10
23	110.5	2.4	469	3	US-08-996-685-10
24	110.5	2.4	593	2	US-08-989-478-2
25	110.5	2.4	593	2	US-08-989-478-8
26	110.5	2.4	593	3	US-08-996-685-2
27	110.5	2.4	593	3	US-08-996-685-8

ALIGNMENTS

RESULT 1

US-09-604-605-2

; Sequence 2, Application US/09604605

; Patent No. 6421613

; GENERAL INFORMATION:

; APPLICANT: Ramgopal Nadimpalli

; APPLICANT: Carl R. Simmons

; TITLE OF INVENTION: Maize Prolifera Gene and Uses Thereof

; FILE REFERENCE: 1138

; CURRENT APPLICATION NUMBER: US/09/604.605

; CURRENT FILING DATE: 2000-06-27

; PRIOR APPLICATION NUMBER: 60/143,222

; PRIOR FILING DATE: 1999-07-09

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 720

; TYPE: PRT

; ORGANISM: Zea mays

US-09-604-605-2

Query Match 16.7%; Score 774; DB 4; Length 720;

Best Local Similarity 32.4%; Pred. No. 4.7e-66;

Matches 200; Conservative 103; Mismatches 241; Indels 74; Gaps 13;

Oy 189 KNFLRTHVDHSHGVFKERISDMCKENRESLVVYEDLAAREHVLAYPL----PEAPAE LL 245

Db 14 KDFLSNFTGPHGEPKYLNLQDVANRKRVAQIELDDLHYKVDDEFLQRVVTENTRRYI 73

Oy 246 QIFDEAALEV-----LAMYPKD-----RITNHIHVR 273

Db 74 GIFAEAMDEIMPEPTTEATYVDEQDILMTQRVDEGAGDGADTDLQKMPPEIKRFFEVY 133

Oy 274 ISHPLVLEEL--SLRQLHNLRTSGVTSCVLPQLSMWYKNCNCF----- 323

Db 134 IKAFSKYPTLTIRQVKASNIQVLKISGIVTRCSDVRLQVAVYTCCEGFEIYQEVTA 193

Oy 324 -VLGPFQSQNQEVKPGSCPECQAGPFVEMETIYONYQIRIQESPGVAARRLPRS 382

Db 194 RVFMPLIECPQSQR-----CKLNKAKGNLILQLRASKFLKFQEVKLQELAEHVPKGIHPR 248

Oy 383 KDAILLADVSCNAGDEIETGIY-HNNYDGSINTANGFPVFATVILANHVAKKDNKA 441

Db 249 LTVHLRGELTRKVPAGDVVENSGLFLPMPYVGFGRAMRAGL-VADTYLEAMSVTHFKKYE 307

Oy 442 VGELTDEVDKMTLSKDDQIGEKIFASIPAGYHEDIKRGPALALFPGGPKNPGGKHK 501

Db 308 EYDLKGEQEQIDRLAEDGDYLSKLARSLAEFIEGHEDVKKALLLLLVGAPKRLADGMK 367

Oy 502 VRGDIINVLLCGDPCTAKSQFLYIEKVSSRAIFTTGGCASAVATYVQRHPVSRWPTLE 561

Sequence 3, Appli
Sequence 281, App
Sequence 21413, A
Sequence 30208, A
Sequence 3159, Ap
Sequence 20, Appl
Sequence 5687, Ap
Sequence 34, Appl
Sequence 52, Appl
Sequence 4, Appli
Sequence 2, Appli
Sequence 21, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 5, Appli
Sequence 57, Appli

ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Tolypocladium niveum
STRAIN: ATCC 34921
US-08-471-119A-2

Query Match 2.7%; Score 125; DB 2; Length 15281;
Best Local Similarity 18.3%; Pred. No. 2.2;
Matches 185; Conservative 124; Mismatches 331; Indels 372; Gaps 47;

QY 7 QRRGRNDPLTSSGRRSRRTDALTSPPGRDLPPFEDESEGLLGT---GPLEEEDEGEEL 63
DB 9156 QYVGLDPAASAA-----IPVNEAVSLPGLACKARVLTALDGLSDKNEIQPEL 9207
QY 64 I-----GD-----GMRDY---RAIPEL-DAYE 82
DB 9208 VVINSVAQYPTSEYLLKVVKAVVEVPSVKRFFGDIRSQALNARDFLAARAVRALGDNAS 9267
QY 83 AEGALDDEDEBELTASRRRAADGPGCTVT--GSPGLGACAVGSCMTAMRRTSRALPAS 140
DB 9268 KEQIR---EKIAELESEEEELLVDPAFFVSLRSLQPLNKHVEV--LPKLMKATNELSSYR 9322
QY 141 AASGAGTEDEGEQMETESTENLEDLKGHSVRWVSMAGPRLEIHHFRKFNFLTHYVDSHG 200
DB 9323 YAAVLHISHNEESQLLIQID-----PTAW-----DFAATQKDSQG 9359
QY 201 HNVFKERISDMCKENRESLV-----VNYEDLAAREHVL----- 233
DB 9360 -----LRNLQGRDDVMIAVGNIPYSKTIYVERHIMNSLDQDHVNSLDGTSWISDARS 9412
QY 234 ----AYFLPEAPAEILQIFDEAALEVVL-----AMYPKYDRITNHHIVRIS--- 275
DB 9413 AAAICTSF--DAPA-LTQLAKEGFRVELSWARQSRONGALDAVFRHLATDANCERSVL 9469
QY 276 -HLPLVEELSLRQLHLNQLIRTSG-----VVTSCGVLPQLSVMKY 316
DB 9470 VHEPTDHQROLRLTLNRPLQRAQSRIIESQVFEALQATALPAYMIPSRILVLPOM----- 9324
QY 317 NCNKNFVLGPFQCSQNOEYKPGSCPCQAGPFEVNMETIYQNTQRIQIESPGKVA 376
DB 9525 -----PTNANGKVDKQLARRAQVAKRAVSA 9552
QY 377 RLPRSKDAIL-----ADLVDSNAGDEIELTGIYHNNDGSLTANGFPVFATVILANH 432
DB 9553 RVAPRNDTEVLCEEAADIL-----GTEV---GITONFED-----MGHSLMATKLAARL 9599
QY 433 VAKKONKVAAGELTDEDVKMITSLSKDOQIGEXIFASIAPIYGHEDIKRGPALALFGE 492
DB 9600 SRLDTRVTKEVFDKPV--LADLAASIEQGSTPHUPIASSVYS-----GPV-----E 9645
QY 493 PKNPGGKHKVRGDIINVLGDPGTAKSQFLKYIEKVSRAIFTTGOCASAVAVTAYVQRH 552
DB 9646 QSYAQGRLEWLDQFNL-----NATWYHMSLAMRLGLPLNMDALDVALRALEQRH 9694
QY 553 PVGREWPLEA-----GALVLADRG---VCLIDEFKMNDORTSIHEAMEQOQTSI-SKAG 604
DB 9695 ETILRT-TFEAQKIDGQVGVHVEAGMKRLKVLDSLK-NEKEHMAVLENEQMRPFTLASERG 9752
QY 605 IVTSLOARCTVIAAANPIGRYDPSLTFSSENVDTLTPILSRFDILCVVRTVDPVQDEML 664
DB 9753 WKHGL-----ARLGP-----TEYLISL----- 9769
QY 665 AREVVGSHVRRHPSN-----KEBEGLANGSAABPAMPNTYGVPELPQEVLPKYYIYA 716
DB 9770 -----VMHHFSDGWSVDILRQELGQFYSALRGDRDLSQ-VKPLPIQ-YEDFAAQW 9819
QY 717 KERV---HPKL-----NQMDQKVAKWYSDLRKESMATGS---IPTVRIHESMSHGGG 764
DB 9820 KEAAQVAEHERQYAWENQLADSTPGELLTDFPRPQFLSGKAGVIPVTIE-----G 9870
QY 765 PRAHPSAGLCDRRRRQHGHPRDAGELHRHTEVORHRSRMRKTFARYLSFRDNNELLFIL 824
DB 9871 P-----VYBKLLKFSKERQVTLFSVL 9891

QY 825 KQLVABQVTVQRNRFGAQOQTIIEVPEKDLVKARQINIHNLISAFYSELFMR 876
DB 9892 --LTAFRATHFR-LTGAEDATIGTP---TANRRNPELEHIGFVNTQCMLR 9937

RESULT 6

US-09-252-991A-24044
Sequence 24044, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 24044
LENGTH: 516
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24044

Query Match 2.6%; Score 121.5; DB 4; Length 516;
Best Local Similarity 23.08; Pred. No. 0.011;
Matches 66; Conservative 43; Mismatches 109; Indels 69; Gaps 11;

QY 504 GDINVLGCGDPGTAKS-----QFLKYIEKVSRAIFTTGOGASAVAVTAYVQR----- 551
DB 229 GAHNLGSGPPGCTKTLASRLPGLPALDEDEALEVAATHSVASHVPLRHPWRPQR 288
QY 552 -HPVSRWTL-----EAGALVLADRGVCLIDFKMNDORTSIHEAMEQOQSISKAG 604
DB 289 HHSASAPALVGGSRPQPGEITLAHQGLVFLDELPEFERKVLVLEPRELSEGEIVARAN 348
QY 605 IVTSLOARCTVIAAANPIGRY--DPS---LTSENV-----LTPEIISREDI-LCVVR 653
DB 349 GRVRFPAFOLVAAAMPNCPGCLGDPGRCRCRTEPQVQRYRGKLSGLLDRIDLHVSVL 408
QY 654 DT-----VDPVQDEMLARFVVGSHVRHHPNKEBEGLANGSAABPAM 695
DB 409 ESTSLQPGHGTATAEVRSERVAARQRLAR-----QGCANAHLDLQAM 452
QY 696 PNTYGVPELPQEVLPKYYIYAKERVHPKLNQMDQD-KVAKMYSDLRK 741
DB 453 HRNCAL-----AADDRRWLEAAGERLELSRALHRLKLVARTLADLER 495

RESULT 7

US-09-107-532A-4537
Sequence 4537, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>

QY 545 VTAY-----VORHPVSEWLEAGALVADRGVCLIDFEDKMNDDQR--TSIHEAM 593
Db 408 RSYLIDULENFEVLEGIYGRW-----LDEIFSIDQFSRMGRKSHQLTEGHOAL 458
QY 594 EQQSISIKAGIVTSIQAQRCIVIAAANPIGGRYDPSLTFSENVD----- 637
Db 459 YKEVIKLNKVNKVLYPE-----QAMNDAHTVVKDYATSFYESLEVPFELPSOLMTORDELE 514
QY 638 ---LTEP-----IISRFIDLCVVRDVTDPVQDEMLARFVVGSHVRHH---PSNKE 681
Db 515 LAGLTEKAEEDQVWNGLIQLDLDLVTFDDQEMTLQOFLDVFDIGLEQLEFVMPQTL 574
QY 682 BEGLANGAA-----EPAMNTYGVVEPLPQEVKLYIYAKERVHPKLN-- 725
Db 575 QVSIGTMDLAKVDNKKHIYMGMDGILPQTVSSSLITDEKKYV---EDNAHVLSPT 631
QY 726 ---QMD-----QDKVAKMYS-----DLRKESMATGSIPIVTRHIESMHHG 763
Db 632 SDILOMDAFVCYIAMTRSQSVTSYSLMNGSGDEKEISPLTOIKELFYDLEILNIQ 691
QY 764 GPRAHPSAGLCRRRRQHGPRDAGELHRRHTEVQRHRSR-----KTFARYLSFR 813
Db 692 LHKAQPLMM-----QHS-----QTKIQLFEVLRGWLHDHEDIDYRWLDAYLAIR 736
QY 814 RDN--NELLFILKQVAVQVYQNRFRGAQODTIEVPEKDLVDKARQINIHNSAF--- 868
Db 737 DDQLNQGLDITTSLSLYDNETVOLNETLSQQ-----LYGKTNASVSREFGYQQC 787
QY 869 ---YDSELFWMKFSDHLKRMILQOF 892
Db 788 PFKYASHGLRLNE-----RTKYLQNF 810

RESULT 9

US-08-755-587-189
; Sequence 189, Application US/08755587
; Patent No. 6045997

GENERAL INFORMATION:

; APPLICANT: Futreal, Phillip A
; APPLICANT: Wooster, Richard F
; APPLICANT: Ashworth, Alan
; APPLICANT: Stratton, Michael R
; TITLE OF INVENTION: Materials and methods relating to the
; TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
; TITLE OF INVENTION: susceptibility gene and uses thereof.
; NUMBER OF SEQUENCES: 222
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell Seltzer Park & Gibson
; STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,587
; FILING DATE: 25-NOV-1996

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9523959.6
; FILING DATE: 23-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9525555.0
; FILING DATE: 14-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9617961.9
; FILING DATE: 28-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenneth D Sibley
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5405-135

; INFORMATION FOR SEQ ID NO: 189:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1589 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

US-08-755-587-189

Query Match 2.6%; Score 118.5; DB 3; Length 1589;
Best Local Similarity 18.1%; Pred. No. 0.17;
Matches 144; Conservative 123; Mismatches 329; Indels 199; Gaps 35;

QY 21 RSSRRTALTSPPGRDLPPFDESEGLLCTGEGPEEEDGELIGDGMERYDRAIPE-LD 79
Db 664 KNTKHEDSYTSQRNNL---ENSDGSMSTSGPYIHKGDSDLPADQSGK---CPSECT 716
QY 80 AYEAGLALDDDEVEELTASRREAADGPGCVTGWPGIACAGVSCMTAMRRTRSLA 139
Db 717 QYARENTQIKENISDLTCLIMKAAE-----TCMKS--SDKQLPS 755
QY 140 SAASGAGTDEGEDEQMIETIE-NLEDLKGHVSREWVSMAGPRLEIHRFKNFRTHVDS 198
Db 756 -----DKMEQNIKEFNISFQTASGKNTRVSKESLNKSNVFNRETDELTVISDS 804
QY 199 HG-----HNVFERISDMCKENRESLVNVED-----LAAREHVLAYFLPEAPAE 246
Db 805 LNSKILHGINKDKMHTSCHKKAISIKKVFEDHPITVTSQLPAAQHP-EYEIESTKEPT 863
QY 247 IFDEAALEVLWAMPKYVDRIITHI---HVRIS-----HLPLVEELSLR----- 290
Db 864 SFHTASGKKVKIMQESLDKVKANLFDQYVVKRTASFSQSGKPLKDSKELFLAYEKI 923
QY 291 LNQLIRTSVVTSCTGVLPQLSMVKY----NCNKNFVLGPPCOS--QNOEVKPGSCPEC 344
Db 924 ASKEEMQNFVSKETEMLPQOYHMYROTENLKTNSGTSSKVQENIENNVKNPRICIC 983
QY 345 QSAGPFVNMETIYQYQIRIOESPCKVAARLPRSKDAILLADLVDSNAGDEIBLT 404
Db 984 QSSYPV---TEDSALAY---TEDSRKTCVRESSLSKGRKWLREQDGKLGTRNTIKIE 1035
QY 405 GIYH--NNYDGLSNTANGFPVFATVLANHVAKKVAVGELTDEDVKMITSLSKDOOIG 463
Db 1036 CKETEDFAGNASEHSLVIRTEIDTNHVE-----NQVSTLLSDPNV- 1079
QY 464 EKIFASIPSTYGHEDIKRGPALALFGGEPKNPG---KHKVRGDIINVLLCGDPGTAKS- 519
Db 1080 --CHSYLSQSFSFCHCD-----DMHNDSGYFLKKNKIDSDVPPDMKNAEGNTISP 1125
QY 520 -----QFLKY-IEKVSSRAITFTTCQGASAVAV-----TAYVQRH 552
Db 1126 RVSATKERNLHPOTINYCVOKLETN---TSPHANKDVAIDPSLDSRNCKVGLSVFITAH 1182
QY 553 PVSREWTLGAGALVADRGVCLIDFEDKMN-DQDRTSIHEAQQSISI-SKAG----- 604
Db 1183 SQETERKE-----IVTDCYKIVQNRQSKPDTCTSCHKVLDDSKDFICSSGDVCLNS 1238
QY 605 -----IVTSLQARCTVIAAANPIG--RYDPSLTFSENVDTPEIISR-FDILC 650
Db 1239 RKDSFCPHNEQILQHNQSMGLKKAATPPVGLTWDTSKSIKREPPQAAHP--SRTYGIFS 1296
QY 651 VVRDTPDVPQD-----EMLARFVVGSHVRHHPKNEEGLANGSAEP 693
Db 1297 TASGKATQVSDASLEKARQVSEMDGAKQLSSMVLGEGNPKPHHSVKRENSVWHSTQGLV 1356
QY 694 AMPNTYGVVEPLPQEV 708
Db 1357 SLP-----KPLPGNV 1366

RESULT 10

US-09-315-793-12

; Sequence 12, Application US/09315793

; Patent No. 6221597

; GENERAL INFORMATION:

RESULT 12
US-09-264-854-5
; Sequence 5, Application us/09264854
; Patent No. 6563025
; GENERAL INFORMATION:
; APPLICANT: SONG, HEE-SOOK
; APPLICANT: BROTHERTON, JEFFREY E.
; APPLICANT: WIDHOLM, JACK M.
; TITLE OF INVENTION: SELECTABLE MARKER AND PROMOTER FOR PLANT TISSUE CULTURE
; TITLE OF INVENTION: TRANSFORMATION

US-08-840-006-6
; Sequence 6, Application US/08840006
; Patent No. 6127520
; GENERAL INFORMATION:
; APPLICANT: Ueda, Tetsufumi
; APPLICANT: Ozkan, Eric D.
; TITLE OF INVENTION: Compositions And Methods For The
; TITLE OF INVENTION: Inhibition Of Neurotransmitter Uptake Of Synaptic Vesicles
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,006
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UM-02786
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1151 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-840-006-6

Query Match 2.5%; Score 116; DB 3; Length 1151;
Best Local Similarity 18.2%; Pred. No. 0.16;
Matches 179; Conservative 156; Mismatches 332; Indels 318; Gaps 49;
QY 35 RDLPPDESEGLTGTEPLEEEDGELIGDMERDYRAIPELDVAEAGLDDDEVE 94
DB 346 RDLTSVWTEKALINAD-ELASDVAGAEAL---LDRHQEHKGIDAHE-DSFKSADESGQ 400
QY 95 ELTASRREADPCGTVTGSPGLGACAGSCMTAMRTSRALPASAAAGAGTEDGEEDE 154
DB 401 ALLAGHYASD-----EVREKLTVLSEERAA-----LELWELRR 435
QY 155 QMTESLENLE---DLKGHSVRWVSMAGPRLRIHHRFNFLRTHVDSHGHNFKERISD 210
DB 436 QQYEQCMQLQFDRDTE--QVDNWS-----KQEAFL-----LNEDLGD 472
QY 211 MCKENRESLVNVED---LAAREHVLAVFLPEAPAEALQIDFAALEVVLAMYPKYDRI 266
DB 473 FL-DSVEALKKHEDFEKLSAQE-----EKTALDEFATKLIQNNHYAMEDV 519
QY 267 TNHIVRHRLPLVEELSLRQ-----LHNLQILRTSGVVTSCGVLPQLSMVKYNCN 319
DB 520 ATRRDALLSRNALHE-RAMRRRAQLADSFHQOFFRDSDELKSWV----- 564
QY 320 KCNFVLGPPCQSQNQEVK-----PGSCP-ECOSAGPFEVNMETIYQNVQRIQOE 369
DB 565 -----NEKMTATDEAYKDPNSNLOGKVQKHOAFEAELS-----ANQSRIDALE 607
QY 370 SPG-----KVAARRLPRSKDAILLADLVDSNAGDEIETGIYHNNYDGLNT 417
DB 608 KAGQKLDVNHAKDEVAAR---MNEVVISLWKLLLEA-----TELKGI-----KLRE 651

QY 418 ANGFPVFATVI-----LANHVAKKDNKVAVAGBELTDEDKMITSLSKDQOIGEKIFA 468
DB 652 ANOQOQFNRRNVEDIELWLVEVEGHASDD-----YKDLTNVQNLQKKHALLBADVA 703
QY 469 STAPSI-----YGHEDI-----KRGPAALALFGGEPKPN---GGKHKVRGDI NVLL 510
DB 704 AHODRDGVTIARQFODAGHFAENIKKKQEAALVARYEALKEPVMARKOKLAUSRLQO 763
QY 511 CGDPGTAKSOFKLYIEKVSRAFTTQGASAVAVTAYVORHP-----VSREWTEAGA 564
DB 764 LFRDVEDEETWIREKEPIAA-----STNRGKDLGIVQNLKKHQAQAEIAGHEPRIKA-- 817
QY 565 LVLADRGVCLIDFDKMNDDRTSIHEAMEQQSISISKAG-----IVTSLQARCTVIAAA 619
DB 818 --VTQKGNAMVEEGHFAAEDVKAKLHELQKWEALKAKASQRQDLEDLSLQAO-QVFADA 874
QY 620 N-----PIGGRYDPSLTFSENVDLTEPIISRFDIL----- 649
DB 875 NEAESWMREKEPIVGTSD---YKDEDSAEALLKKHHEALMSDLSAYGSSIQALREQAQS 930
QY 650 CVVRDTPVDPVDEMLARFVGVSHVRHPSNKE---BEG-----LANG-----SAAE 692
DB 931 C--RQQVAPDDETGEKELVLYDYQEKSPREVTMKKGDLITLLNSTNKKDWKVEVNDRO 988
QY 693 PAMPNTYGEPL--POEVLKYYIIYAKERYVHPKLNQMD-QDKVAKMYSDLRKESMATGSI 749
DB 989 GFVPAAY-VKKLDPAQSASRENLEEQSGIALRQEQIDNQTITK-----EAGSV 1037
QY 750 PITVRHIESMSHGCGPRAHPAGLCDRRRRHOHGHPRDAGELHHRHTEVORHRSRKTFARY 809
DB 1038 SLRMKQVEELYHS-----LLELGE-----KRGKMLEKSKCKF 1069
QY 810 LSFRDNNELLFILQO---LVAEQVTYQNRNFGAODTIEVPEKDLVDKAROINIHLS 866
DB 1070 MLF-REANELOQWINEKEAALTSEEV-----GADLEQVEVLQKKFDDQKDLKAN--- 1118
QY 867 AFYDSELFMRNKFSDLRKRMILQO 891
DB 1119 ---ESRLKDINKVAEDLESEGLMAE 1140
RESULT 15
US-08-840-006-5
; Sequence 5, Application US/08840006
; Patent No. 6127520
; GENERAL INFORMATION:
; APPLICANT: Ueda, Tetsufumi
; APPLICANT: Ozkan, Eric D.
; TITLE OF INVENTION: Compositions And Methods For The
; TITLE OF INVENTION: Inhibition Of Neurotransmitter Uptake Of Synaptic Vesicles
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,006
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UM-02786
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2003, 13:13:55 ; Search time 70 Seconds
(without alignments)
1859.341 Million cell updates/sec

Title: CAA47749
Perfect score: 4630
Sequence: 1 MASSPAQRGRNDPLTSPG.....LFRMKNFSDLRKRMILQOF 892

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4078.5	88.1	904	15	US-09-964-899-49
2	1872	40.4	864	10	US-10-128-714-8071
3	1840.5	39.8	785	15	US-10-128-714-3071
4	1073.5	23.2	268	10	US-09-925-300-1461
5	891	19.2	1023	15	US-10-128-714-8558
6	887	19.2	803	15	US-10-128-714-8363
7	882.5	19.1	831	15	US-10-128-714-3363
8	860	18.6	874	15	US-10-128-714-3558
9	835	18.0	821	15	US-10-171-311-129
10	835	18.0	901	11	US-09-828-062-8
11	810	17.5	718	15	US-10-128-714-3399
12	804	17.4	713	15	US-10-128-714-8399
13	783	16.9	724	10	US-09-925-300-1053
14	778	16.8	766	9	US-09-925-301-1382
15	776	16.8	921	15	US-10-128-714-3372

16	774	16.7	720	15	US-10-101-080-2	Sequence 2, Appli
17	763	16.5	819	9	US-09-833-790-425	Sequence 425, App
18	754	16.3	957	15	US-10-128-714-8372	Sequence 8372, Ap
19	753.5	16.3	814	15	US-10-128-714-8100	Sequence 8100, Ap
20	732.5	15.8	632	15	US-10-128-714-3100	Sequence 3100, Ap
21	451	9.7	94	9	US-09-925-299-1224	Sequence 1224, Ap
22	451	9.7	94	11	US-09-925-299-1224	Sequence 1224, Ap
23	451	9.7	113	15	US-10-106-698-5543	Sequence 5543, Ap
24	236	5.1	69	9	US-09-864-761-46505	Sequence 46505, A
25	201	4.3	171	15	US-10-106-698-5506	Sequence 5506, Ap
26	193.5	4.2	63	9	US-09-864-761-45117	Sequence 45117, A
27	186	4.0	66	9	US-09-925-302-730	Sequence 730, App
28	154	3.3	138	10	US-09-764-868-939	Sequence 939, App
29	139	3.0	178	15	US-10-106-698-5566	Sequence 5566, Ap
30	132.5	2.9	367	10	US-09-738-626-5688	Sequence 5688, Ap
31	129.5	2.8	872	14	US-10-047-260-38	Sequence 38, Appl
32	127	2.7	507	10	US-09-738-626-5725	Sequence 5725, Ap
33	124	2.7	867	14	US-10-007-693-109	Sequence 109, App
34	121	2.6	1177	15	US-10-128-714-3493	Sequence 3493, Ap
35	121	2.6	1179	15	US-10-128-714-8493	Sequence 8493, Ap
36	120.5	2.6	856	9	US-09-815-242-11310	Sequence 11310, A
37	120.5	2.6	1324	12	US-10-205-219-113	Sequence 113, App
38	119.5	2.6	541	15	US-10-156-761-10169	Sequence 10169, A
39	117.5	2.5	856	9	US-09-815-242-11489	Sequence 11489, A
40	117	2.5	616	15	US-10-138-927-103	Sequence 103, App
41	115	2.5	1481	15	US-10-331-061-70	Sequence 70, Appl
42	115	2.5	2910	14	US-10-124-800-2	Sequence 2, Appli
43	113.5	2.5	882	9	US-09-815-242-10668	Sequence 10668, A
44	113.5	2.5	6304	15	US-10-147-026-16	Sequence 16, Appl
45	113	2.4	2472	9	US-09-815-242-5064	Sequence 5064, Ap

ALIGNMENTS

RESULT 1

US-09-964-899-49
; Sequence 49, Application US/09964899
; Patent No. US20020174446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; FILE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964.899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236.893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298.309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-964-899-49

Query Match	88.1%	Score	4078.5	DB	10	Length	904
Best Local Similarity	89.8%	Pred. No.	0				
Matches	817	Conservative	8	Mismatches	52	Indels	33
Gaps	6						
Qy	1	MASSPAQRGRNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGTEGPLEEEDG	60				
Db	10	MASSPAQRGRNDPLTSSPGRRRTDALTSSPGRDLPPEDESEGLLTGTEGPLEEEDG	69				
Qy	61	EELIGDGMERYRAIPELDADYEAGLDDDEVELTASRREA-----DGPCTGVTG	113				
Db	70	EELIGDGMERYRAIPELDADYEAGLDDDEVELTASRREAARMRQRDEAGR---	126				
Qy	114	SWFGLGACAVGSCMTA-----MRTSRALPASAGAGTDEGEDEQMIENLED	165				
Db	127	---GLGMRRLGLLYDSDEDEERPKRRQVERA-----TEDGEDEMIENLED	176				

Db 763 LRRKYLYAREKCHPKLYQIDQDKVARLFADMRRESLATGAYPTVVRHLEAI 814

RESULT 3

US-10-128-714-3071

; Sequence 3071, Application US/10128714

; Publication No. US20030119013A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Bo

; APPLICANT: Hu, Weng

; APPLICANT: Tishkoff, Daniel

; APPLICANT: Zamudio, Carlos

; APPLICANT: Broshkin, Alexey M

; APPLICANT: Lemieux, Sebastien M

; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

; TITLE OF INVENTION: Methods of Use

; FILE REFERENCE: 10182-018-999

; CURRENT APPLICATION NUMBER: US/10/128,714

; PRIOR FILING DATE: 2002-04-23

; PRIOR APPLICATION NUMBER: US 60/285,697

; PRIOR FILING DATE: 2001-04-23

; PRIOR APPLICATION NUMBER: US 60/287,066

; PRIOR FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: US 60/295,890

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/303,899

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/316,362

; PRIOR FILING DATE: 2001-08-31

; NUMBER OF SEQ ID NOS: 8603

; SOFTWARE: Patent version 3.1

; SEQ ID NO 3071

; LENGTH: 785

; TYPE: PRT

; ORGANISM: Aspergillus fumigatus

US-10-128-714-3071

Query Match 39.8%; Score 1840.5; DB 15; Length 785;

Best Local Similarity 49.0%; Pred. No. 2.2e-152;

Matches 388; Conservative 102; Mismatches 207; Indels 95; Gaps 10;

QY 33 PGRDLPPF-----EDESGLLGTEGLE-----EEDGEEELIGDMERDVRAPELDA 80

DB 5 PPSPPMLPDEDDIETDEEAEELGIDIDDEMAEDGIDLFNGDFERYDVGNG-DR 63

QY 81 YEAEGALDDEVEEELTASREAAAGPCGTVGSPGLGACAVGSCWTAMRRTRSLPAS 140

DB 64 YGEYID-DGGHEELDIAIRQLRLN-----RRDRELEERR 101

QY 141 AASGAGTEGEE-----DEQMTIE-SIENLEDLKHSVRE 173

DB 102 RMPAFLQDDDEGDMDLTRQPRRRHHYDEDEDIEMGDDAMEELSLEELADVKAANITD 161

QY 174 WYSMAGPRLEIHRFNKFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAREHVL 233

DB 162 WLQPOVLRSIYREKFAFTTEFTDQAGQSVYGHKIKTLGEVNSASLEVSVHLSSSTAAL 221

QY 234 AYFLPEAPAEELLIQFDEAALEVVLAMYPKYDRITNHIHVRISHLPLVEELRSRLHLNQ 293

DB 222 SYFLANEPEVLKVFQDAVDVTLFHYPOQDIHKHVRISDLPVLYTLRQROHLNC 281

QY 294 LIRTSQVVTSCITGLVLPQLSWVYKNCNFKVLPFCQSQNQYKPGSCPECQSAQPFVFN 353

DB 282 LRVSGVVRTTGTGVPQLAYVMFICQKCNITLGPFOQEAQAEVKSICYCQNSQSGPFTIN 341

QY 354 MEETIYQNTIRIQSPQKVARRLPRSKDAILLADLVDSNAGDEIELTGYHNNYDG 413

DB 342 SEKTVYRNTQKTLQESPSGVPAGRLPRQREVLLADLDSAKPGDEIEVTGYRNSYDA 401

QY 414 SLNTANGPPEVATVILANHVAKONKVAVGELTDEDVKMTISLSDQOIGCEKIFASIAPS 473

DB 402 QLNKNGKFPVFATIEIANHVSHQDLGFLHTEDEQIRALSRDPDIVDKIVRSIAAPS 461

QY 474 IYXHEDIKRGPALALFGEKPNPGGKHKVGRGINVLLCGDPGTAQSOFLKYIEKVSRAI 533

DB 462 IYGHQDVKTAVALSFLPGVSKAQAQKMAIRGIDINVLLGLDGPCTAKSOVLKYVEKTAHRAV 521

QY 534 FTTGGASAVAVTAYVORHPVSRWTLERAGALVLAIRGVCCLIDEFKMMNDODRTSHEAM 593

DB 522 FATGGASAVGLTASVRRDPLTSEWTLLEGALVLAIRGVCCLIDEFKMMNDODRTSHEAM 581

QY 594 EEOISISKAGIVTSLQARCTVIAAANPIGGRYDPSLTFSENVDLTERPIISRFDILCVVR 653

DB 582 EEOISISKAGIVTTLQARCAVAAANPIGGRYNSIPIFSQNVLELTERPILSRFDILCVVR 641

QY 654 DTVPVQDEMLARFVVGSHVRHHSNK-----EEDG----- 684

DB 642 DTVPDNEDELANFVIESHHRANPTRPLRDQGNLVDSEGNRIDESGYRLDKHGNRLPPT 701

QY 685 ---LANGSAEAPAMPNTYGVPELPQEVLLKYYIYAKERVHPKLNQMDQDKVAKMYSDLRK 741

DB 702 PEEIAKREAAQKAEKEEKEGE-IPQELLRKYYILYAREKCHPKLYQIDQDKVARLFADMR 760

QY 742 ESMATGSIPITV 753

DB 761 ESLATGAYPITV 772

RESULT 4

US-09-925-300-1461

; Sequence 1461, Application US/09925300

; Patent No. US20020151681A1

; GENERAL INFORMATION:

; APPLICANT: Craig Rosen,

; APPLICANT: Steve Ruben,

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA101

; CURRENT APPLICATION NUMBER: US/09/925,300

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05988

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 1890

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 1461

; LENGTH: 268

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-925-300-1461

Query Match 23.2%; Score 1073.5; DB 10; Length 268;

Best Local Similarity 91.7%; Pred. No. 9e-86;

Matches 211; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

QY 1 MASSPAQRRRGNDPLTSSPGRSSRRRTDALTSSPGRDLPPFEDESEGLLGTEGLEEEDG 60

DB 30 MASSPAQRRRGNDPLTSSPGRSSRRRTDALTSSPGRDLPPFEDESEGLLGTEGLEEEDG 89

QY 61 EELIGDMERDVRAPELDAYEAGLALDDEVEELTASREAAAGPCGTVTGSWPGL-G 119

DB 90 EELIGDMERDVRAPELDAYEAGLALDDEVEELTASREAAAGPCGTVTGSWPGL-G 149

QY 120 ACAGSCMTAMRRTRSLPASASAGAGTEDGEDEQMIESIENLEDLKHSVREWYSMAG 179

DB 150 ACAGSCMTAMRRTRSLPASASAGAGTEDGEDEQMIESIENLEDLKHSVREWYSMAG 209

QY 180 PRLEIHRHFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDLAAR 229

DB 210 PRLEIHRHFKNFLRTHVDSHGHNVFKERISDMCKENRESLVVNYEDTGSQ 259

RESULT 5

US-10-128-714-8558

; Sequence 8558, Application US/10128714

; Publication No. US20030119013A1

```
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroszhkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8558
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8558

Query Match      19.2%; Score 891; DB 15; Length 1023;
Best Local Similarity 28.7%; Pred. No. 9.3e-69;
Matches 268; Conservative 140; Mismatches 304; Indels 222; Gaps 29;

Qy 1 MASSPAQRRC--NDPLTSPGSSRRRTDALTSPGRLPPFDE-SEGILGTGEPLEE 56
Db 1 MSSPASSRRGRPAKDATSPSTRSRQLQTSPP--TPRADEQSOATPPASRRRLR- 56
Qy 57 EEDGEELIGDMERYAIPELDAYEAGLALDDDEVEELTASRRAAD-----GP 107
Db 57 ---GEEAVPSSPFFQSSPSKADSAE---TPDVRDEPSSPRESSTWDEGDRTPGN 110
Qy 108 CGVTGSGWPGAGACAVGSCWTAMRR--TRSALPASA-----141
Db 111 APTMRDSSP-IRYMSSSSPTRAQNRQSRSDIPSSSGLFVSSRPSIESNRAVSRSDLH 169
Qy 142 -----ASGAGTEDGEDEQMIESIENL-----EDLKGHSVRE-WVS 176
Db 170 SGGFLSSPNRRRRVFDVANGMPATDG--DPRSDATFSNIHPDTSEAEALGGSSTRVW-- 225
Qy 177 MAGPRLETH---HRFKNFL-----RTHVDSHGHNVFKERISDMCKENRE----- 217
Db 226 --GTNISIQDSMSAFKNFLYNFQTKYRLWARGATEDETRINGDSAEEREYISMLSTMRLQ 283
Qy 218 ---SLVNVYEDLAAREHVLAFF--LPAPAEALQIFDEAALEVVLAM----- 259
Db 284 GVTSLNLDKAKLNPSTKLWHLQAHAYPQBIIPMBQTVKDVNMVLAIKEMERLRAQNQ 343
Qy 260 -----YPKYDRITNIHVRISHLP-LVEE-----LRLSL 286
Db 344 RNQHNHGLSGPAVPSSDALSETGRMPQNPQIPDLVGEVETKAFKVLFPGLDSTVNMRLD 403
Qy 287 RQLHNLQIRTSQVVTCTGVLPLPSLMVKYCNCKNFVLGPFQSQNQEVKPGSCPE--C 344
Db 404 DPADMKLVSLKGLVIRTPIDPKKEAFFRCQVCNHGV-QVQIDRCKIAEPTCEPRVC 462
Qy 345 QSAGPFVNNMEETIYQYRIRIQESPGKVAARLPRSKDAILLADVDSNAGDEIELT 404
Db 463 KERNSMQLIHRNRCVFADKQVLIKLOETPDSPDGQTPHSVSLCVYDELVDVCKAGDRVEVT 522
Qy 405 GIYHNNYDGSINTANGFPV-----FATVILANHVAKDNK-----439
Db 523 GIFRCN-----PVRVNPQRQTKSLFKTYIDVLHVQKIDRKKLGLIDVSTIEQEL 571
```

```
Qy 440 --VAVGE-----LTDEDVKMITSLSKDDQIGEKIFASIAPIASYXHEDIKRGPALALFGG 491
Db 572 SEQAAGDAEQTRRLTAEEREXIKRTATRPDLVYELLSRLAPSIVEMDDVKKGILLQLFGG 631
Qy 492 EPK--NPGGKHKVRGDIINVLLCGOPGTAKSKOFLKYEKVSRAITFTGQGGASAVATYAV 549
Db 632 TNKTFQKGNPRYRGDINILCGDPSTSKSOLLRYVHKIAPRGVYVTSKGSSAVGLTAYV 691
Qy 550 QRHPVSREWTLAEGALVLADRGVCLIDFEDKMNDDRTSIHEAMEQQOSISIKAGIVTSL 609
Db 692 TRDPETROMVLESALVLSGGGICCIDFDMKNSTESVLEHVEWQQTVSTAKAGIITLL 751
Qy 610 QARCTVIAAANPIGGRYDPDLTFSENVDLTPIISRFEDILCVVRDTPVQDEMLAREVW 669
Db 752 NARTSILASANPIGSRYNPPLVPQNIIDLPTLLSREDLVYLVLDVRDEQEDRLAKHLV 811
Qy 670 GSHVRRHPSNKEEGLANGSAAPMPNTYGVPELPQEVLPQVLLKYYIYAKERVHPKLNQMDQ 729
Db 812 NWYLEDREHAEEQ-----EILPIEFLTAYITYAKTKVHPVLTPAAG 853
Qy 730 DKVAKMYSDLRKES----MATGSIPITVVRHTESM 759
Db 854 KALSDAYVMNRKLGDDIRSSDRRITATTRQLES 887

RESULT 6
US-10-128-714-8363
; Sequence 8363, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroszhkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8363
; LENGTH: 903
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8363

Query Match      19.2%; Score 887; DB 15; Length 903;
Best Local Similarity 33.3%; Pred. No. 1.7e-68;
Matches 230; Conservative 109; Mismatches 252; Indels 100; Gaps 13;

Qy 221 VNYEDLAAREHVLAFFLPAPAEALLOIFDEAALEVVLAMYPKYDRIT-NHIVRISHLP 279
Db 66 VSDIEIRAHNRELADGLLTSPFDYSLAFDRALKVEIKTLPNRPRESRETTADDVNYCAVYGA 125
Qy 280 VEEL-----RSLRQLRHLNQLIRTSQVVTCTGVLPLQ-LSMVKYCNCKNFVLGPFQCS--- 331
Db 126 FGEFSCNPRTLGSSTHLNRMISLEGIVTKCSLVRPKIIQSVHYHNERKORFVARRYRQDTMT 185
Qy 332 ----QNOEVKPGSCPECOSAGPFVNNMEETIYQYRIRIQESPGKVAARLPRSKDAIL 387
```



```

; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3558
; LENGTH: 874
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3558

Query Match      18.6%; Score 860; DB 15; Length 874;
Best Local Similarity 29.6%; Pred. No. 3.8e-66;
Matches 251; Conservative 127; Mismatches 269; Indels 200; Gaps 25;

Qy  2  ASSPA-----QRRRGNDPLTSS-----PGRS-SRRTD-----ALTSFGRDLPFF 40
Db  3  SSSPTAQNQRSSRDIPSSSGLFVSSRPSIESNRAVSRSDLSHSGGLFSSPNRRRVF 62
Qy  41  EDESEGLLGTGPLEEEDEGEELIGDMERYDRAIPELDAYEAGLALDEDEVEELTASR 100
Db  63  VD-ANGMPATDGD-----PRSDA-TFSNIHPDTSEAEALGGSS 98
Qy  101  REAADPGCVNVTGWPGLGACAVGSCWTAMR-----RTRSALPASAASGAGTGEDEE 154
Db  99  TRVINGTNSIQDS-----MSAFKNFLYFQTKYRLWAGA-----TED--EYR 140
Qy  155  QMIESIENLEDLKGHSVREKVMAGPRLEIHHRFKNFLRTHVDSHGHNVKERISDMCKE 214
Db  141  IMGDSAE-----REVISM-----LSTMROL 161
Qy  215  NRESLVVNYEDLAAREHVLAYF--LPEAPAEILQIFDEAALEVVLAMYK-----YD 264
Db  162  GVTSNLDAKNLAKAYPSTLKLWHLHAYPQEIILPMDQTVKQVMVELAIKEMERLRAQN 221
Qy  265  RITNH-----IHVRISHPLVEELSRQLHNLQIRTSQVVTSCVGLPOLSM 313
Db  222  RQNHNRGLSSVETKAFKVLPGLDSTVNMRLDPADMDKLVSIKGLVITRTIIDMK 281
Qy  314  VKYCNKCNFVLGPFQSQNQEVKPGSCPE--CQSAGPFVNMEEIYQNYQIRIOESP 371
Db  282  AFFRCQVCNHGV-QVDIDRGKIAEPTCEPRVCKERNQMOLIHNRVFAEDKQVILQETP 340
Qy  372  GKVAARLRPSKDAILLADLVDSNAGDETELGIYHNNYDGSINTANGFPV-----423
Db  341  DSIPDGTQPHSVSLCVYDELVDVCKAGDRVEVTGIFRCN-----PVRVNPQR 389
Qy  424  ----FATVILANHVAKKDNK-----VAVGE-----LTDDEVDKMTISLSK 458
Db  390  QKSUFKTYIDLVHVKIDRKLKGLIDYSTIBQELSEQAGDAEQTRRLTAEEBKIRAT 449
Qy  459  DQOIGEKIFASIAPSYIGHEDIKRGPALALFGGEPK--NPGGKHKVGDNVLLCGDPGT 516
Db  450  RPDLYELLRSLASPSIYEMDVKKILLQLFGGNTKTFQKGNPRYRGDINILCGDPST 509
Qy  517  AKSOFKLYIEKVSRAIFTTGOGASAVATYVQRHPVSRWETLEAGALVLDGRVCLID 576
Db  510  SKQLRLYHVKIAPRGVITSGKSSAVGLTAYVTROPETQMWLESALVLSGGICCID 569
Qy  577  EFKMNDQDRTSITHAMEQOSISKAGIVTSLQARCTVIAAANPIGGRYDPSLTFSEN 636
Db  570  EFKMNESTRSLVHEVMEQTVSIKAGIITLNRATSIILASANPISGRYNPLVPQNI 629
Qy  637  DLTEPIISREDILCVVRDVTDPQDEMLARFVGVSHVRHHPSNKEEBGLANGSAAEPAMP 696
Db  630  DLPPTLLSRDLVVLDRYDEQEDRLAKHLNMVYLEDREPEHAAEQ-----676
Qy  697  NTYGVPELPQEVLLKXYIAKERVHPKLNQMDQDKVAKYISDLRKES---MATGSIPIT 752
```

```

Db  677  -----EILPIEFLTAITYAKTKVHPVLTAPAAGKALSDAVYNNMRKLCDDIRSSDRRTAT 731
Qy  753  VRHIESM 759
Db  732  TROLES 738

RESULT 9
US-10-171-311-129
; Sequence 129, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-129

Query Match      18.0%; Score 835; DB 15; Length 821;
Best Local Similarity 27.5%; Pred. No. 5.4e-64;
Matches 236; Conservative 155; Mismatches 284; Indels 182; Gaps 22;

Qy  137  LPSAASGAGTGEDEGEDEQMIENLEDLKGHSVREKVMAGPRLEIHHRFKNFLRTHV 196
Db  3  LAAAEPPGAGSQHLEVRDEVAEKQKL-----FLDFLEEFQ 38
Qy  197  DSHGHNVRERISDMCKENRESLVVNYEDLAAREHVLAYFLPEA-----PAELQIED 249
Db  39  SSDGEIKYLOLAABELIRPERNTLVVSFVDLEQFNQQLSTTIOEEFYRVYPYLCALATFV 98
Qy  250  EAALVVLAMYPKYDRITNHIHVRISHPLVEELSRQLHNLQIRTSQVVTSCVGLP 309
Db  99  KDRKEPLA-----KDFYVAFQDLPTRHKIRELTSSRIGLLTRISQGVVTRHPV 149
Qy  310  QLSMWKYNCKNFVGLGPFQSQNQEVKPGSC--PECQSAGPFVNMEEIYQNYQIRI 367
Db  150  ELVSGTFLCQDCQTVIRD-VEQQFKYTPQNPICRNPVCANRRRFLDNTKSRFVDFQK 208
Qy  368  QESPGKVAARLRPSKDAILLADLVDSNAGDETELGIY-----HNNYDG 413
Db  209  QETQAEULPRGSIPIRSLVILRAEAVESAQAGDKCDFTGTLIVVPDVSKLSTPGARAE 268
Qy  414  SLANTANGFPV-----FATVILANHVAKKDNKVAAGELTDEVD-----450
Db  269  RVSCVDGYETEGIRLGRALGVRLDSYRLVFLACCAVPTNPRFGGKELRDEEQTAE 328
Qy  451  -----KMTYSLSKDQOIGEKIFASIAPSYIGHEDIKRGPALALFGGEPKPNKGK 504
Db  329  MTKWEKVEKVEFMSQDKNLYNLCTSLPTTHGNDVVRGVLMLFGVGPVTKTGGTSL 388
Qy  505  DINVLGCDPGTAKSQFLKYIEKVSSRAIFTTGGASAVATYVQRHPVSRWETLEAGA 564
```

Db 389 DINYCIYDGPSTAKSOFKLVHEEFSRAVYTSKGASAAAGLTAAVVRDEESHEVIEBAGA 448
Qy 565 LVLDARGVCLIDEDKNDODRTSIHEAMEOOSTISIKAGIVTSIQARCTVIAAANPTGG 624
Db 449 LMLADNGVCCIDEFDKMDVDRQVAIHEAMEOQTSITIKAGVKATLNARTSILAAANPLSG 508
Qy 625 RYDPSLTSFENVOLTEPIISRFILCVVRDTPVQVDEMALARFVVGSHVRHHPNKEEG 684
Db 509 HYDRSKSLQONILNAPIMSREFDLFFILVDCNEVTDYAIARRVDLH----- 556
Qy 685 LANGSAEPAMPNTYGVPEPLQEVVKYIYAKERVHPKLNQMDQKVAKMYSILRK--- 741
Db 557 ---SRIEESDRYSLDD-----IRRYLLFARO-FPKPKISKESEDFIVEQYKHLRORDG 606
Qy 742 ESMATGSIPIVTRHIESMNH--GGGPRHPSAGLDCRRRROH----- 781
Db 607 SGVTKSSWRITVROLESIMIRLSEAMRH-----CCDEVQPKHVKEATFLLNKSIIIVETP 662
Qy 782 -----GH---PRDAGELHRRHTEVQRHRSRMRKTFAR--YLSFRR 814
Db 663 DVNLQDEEIQMEVDEGAGGINGHADSPAPVNGINGYNEDINOESAPKASRLGFSEYCR 722
Qy 815 DNNELLFI-----UKQLVAEQVTVORNRFGAQODTIEVPEKDLVDKARQIN-- 861
Db 723 ISNLIVLHLRKVEEEDDESALKR--SELVNNYLKEIESEIDS---BEELINKKRIIEKV 776
Qy 862 IHNLSAFVDSLEFRMKN 878
Db 777 IHRUT-HYDHLVLIETQ 792

RESULT 10

US-09-828-062-8
; Sequence 8, Application US/09828062
; Publication No. US20030097675A1
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: BOHNERT, HANS J.
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUYING
; APPLICANT: SARRIA-MILLAN, RODRIGO
; TITLE OF INVENTION: CELL CYCLE STRESS-RELATED PROTEINS AND METHODS OF USE
; FILE REFERENCE: 16313-0031
; CURRENT APPLICATION NUMBER: US/09/828,062
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/196,001
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 901
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-828-062-8

Query Match 18.0%; Score 835; DB 11; Length 901;
Best Local Similarity 27.5%; Pred. No. 6.3e-64;
Matches 247; Conservative 151; Mismatches 348; Indels 152; Gaps 24;
Qy 2 AASP-AQRRNGNDPLTSPGSRSSRTDAL--TSSPGRDL-----PPFEDESEGLGT-- 50
Db 38 AASPVAGRAVROPTSAVRGRGETDSARRRRSRSLGNSVYSSPDAGTGPPTGPV 97
Qy 51 EGPLEEEDGELIGDGMERYRAIPELDAYAEGLALDDEVEELTASRREAADGPGCT 110
Db 98 ATPVATPVGTPM---GTPSPHRTGPQY-----KORSE----- 127
Qy 111 VTGSPWPLGACAVGSCWTAMRTSALPASASGAGTGDGEDSOMIESIENLEDKHS 170
Db 128 -----LG--SOGKPLHRRRSQSREFGHRP---SREPSADGRPSAEPPDITLGGEY 175

Qy 171 VREWVSMAGPRL---EIHHRFKNFRTHVDHSHGHNV---FKERISDMCKENRESLWNVE 224
Db 176 AYVW---GTNNIPDVLRAIRRELHNY-RSSADHLSKYIQTTEETVEREEDTLNIDMS 230
Qy 225 DLAREHVLAYFLPEAPAELOQIFDEARALEVVLAMYPKYDRITHIHHVIRSHLPLVEELR 284
Db 231 DIYDHDPDLAKIVRYPLDIIPLDTECEQEVATSLPTFEK---HIEARPFLNKASVHMR 287
Qy 285 SLRQHLNOLLRTSGVVTCTGVLGPOLSMWYKNCNKFVLGPFQCSQONQVKGSC--P 342
Db 288 ELNFSDDIKLVSGVMVIRCSIIPEIKGAFKCLVCOHSPPLTVVYKGRVEEPTCRCKP 347
Qy 343 ECQSAGPEVMEETIYQYQIRIQESPGKVAARRLPKSKDAILLADLVDSNAGDEIE 402
Db 348 ECAARNAMSLHNRCTFANKQIVRLQETPDALPEGETPHTVSMCLYNMTVDVAVKPGDRIE 407
Qy 403 LTGIYHN---NYDGLNLTANGFPFVATVILANHVAKKDNKVAVGELTDED----- 449
Db 408 VTGVFKAMAVRVGNQRTLRA--LYKTYIDCVHYKSKDR---GRLOTEDPEMDKENDM 461
Qy 450 -----VKMITSLSKDOQICEKIFASTAPSIYGHEDIKRGPALALFGGE 492
Db 462 YAGHESDTSAAANEAKIOKLKELSGIPGIDYDRLSRSLAPSIWELEDIKGLCOLFGCK 521
Qy 493 KPNPGGHKVRGDIINLLCGDPGTAKSOFKLYIEKVSSRAITFTTGGQASAVATAYVQRH 552
Db 522 AKKIPSGASPRGDIINLLVGGDGTSGSKQLLYVHKIAPRGIVTSGRGSASVGLTAYVTKD 581
Qy 553 PVSREWLEAGALVLADRGVCLIDFEDKMNODRTSIIHEAMEQOSISIKAGIVTSLOAR 612
Db 582 PETRETVLESAGLVLSDRGICCIDFEDKMSDNARSMLEHVMEQQTVSVAKGGLIASLNAR 641
Qy 613 CTVIAAANPIGGRYDPSLTFSENVDLTPETIISRFILCVVRDTPVQVDEMALARFVVGSH 672
Db 642 TSVLACANPSGRNARLUSVDINQLPPTLSRLSDIYLMLDKPDQEQNDRRLARHLVALH 701
Qy 673 VRHPSNKEEGLANGSAAEPAMPNTYGVPEPLQEVVKKYIYAKERVHPKLNQMDQKV 732
Db 702 YENYEVSKQ-----DALDLQTLTAYIYARQHVHPTLSDEAAEDL 741
Qy 733 AKMYSDLRKESMATGS-----IPIVRHIESHGGGPRHPSAGLDCRRRQHGHPRDAG 788
Db 742 INGVEMRQKGNFPGSKKVTATPRQLES-----IRISEALARMRFSEVVEKVDAA 794
Qy 789 ELHRHTEVQRHRSRMRKTFARVLS-----FRDNNELLLFTLKOLVABQVT 833
Db 795 EAVRLDVALQOQSATDHATGTIDMDLITTVGSASERIRAN---LLAALRELITADKIS 849

RESULT 11

US-10-128-714-3399
; Sequence 3399, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Ershkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09


```

QY      471 APSIYGHEDIKRGPALALFGGPKNPGGKHKVRGDINVLLCGDGPCTAKSQFLKYIEKVSS   530
DB      388 APEITYGHEDVKKALLLLLLLVGGVDQSPRGM-KIRGNINICLMGDPGVAKSOLLSYIDRLAP   446
QY      531 RAIFTTGCASAVANTAYVQRHPVSREWTLEAGALVLADRGVCLIDFDFKMNDQDRSTH   590
DB      447 RSQYTTGRSGSSVGLTAALVDRSVSGELTLEGALVLADQGVCCEIDFDFKMAEADRTAHH   506
QY      591 EAMEQOSISIKAGIVTSLOARCTVTIAAANPIIGRYDDPSLTFSENVDLTEPIISRFDILC   650
DB      507 EVMEQQTISIAGKIITLNARCSTIIAAANPAYGRYNRRSRLEQNTQLPAALLSRFDLLW   566
QY      651 VVRDTPDVQDEMLARFV--VGSHVRHPPSNKEEEGLANGSAAEAPAMPNTYGYVEPLPOEV   708
DB      567 LIQDRPDNRDLRLAQHIITYVVHQHSRQPFSQ-----FEPLDMKL    605
QY      709 LKYYIIYAKERVHPKLNQMDDQKVAKMYSDLRKESMAT   746
DB      606 MRRYIAMCREK-QPMYPESPESLADYITAAYVEMRREAWAS   642

RESULT 15
US-10-128-714-3372
; Sequence 3372, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3372
; LENGTH: 921
; TYPE: PR1
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3372
```

Search completed: September 12, 2003, 13:23:14
Job time : 75 secs